Title
Structure and function at a replication origin region within the tetrahymena thermophila rDNA macronuclear chromosome

Author
Gallagher, Renata Constance

Publication Date
1996

Peer reviewed|Thesis/dissertation

Structure and Function at a Replication Origin Region Within the *Tetrahymena thermophila* rDNA Macronuclear Chromosome

by

Renata Constance Gallagher

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Department of Biochemistry

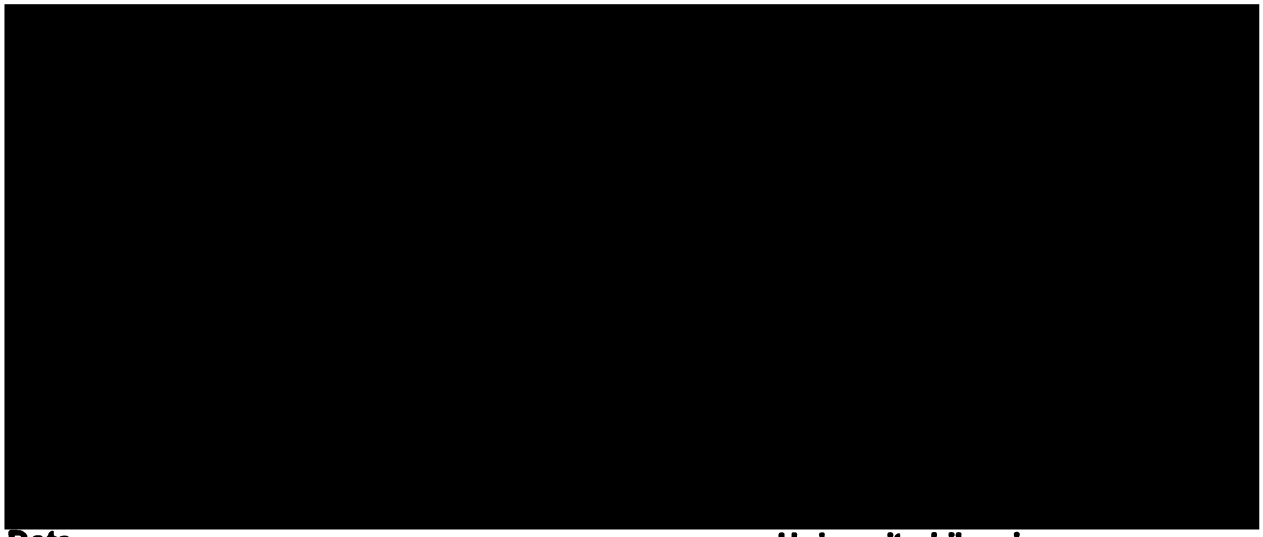
in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco



Date **University Librarian**

Degree Conferred: . . . . . .

Copyright Page

This thesis is dedicated to the memory of my

grandfathers:

Francis Patrick Gallagher

Otto Ewald Dohrenwend

and to my friend,

Michael Komaromy

ACKNOWLEDGEMENTS

Those who think that a scientist leads an isolated, lonely life have never met one. As I begin to thank the many people who have helped me get to... and through... graduate school I am flooded with names and memories. I will do my best to acknowledge the most important and most recent influences on my personal and scientific lives. I know that I will leave out some, and I thank everyone I mention here, and numerous others for their influence and support.

UCSF has been a wonderful place in which to spend more than a decade as a student. First and foremost I owe a great debt of thanks to Nancy Craig for allowing me to spend my year off from medical school in her laboratory. I viewed that experience as my last chance to "do science" and I have been doing it ever since. Nancy provided me with a fun, important, and doable project. She supported me in my decision to take a graduate class while working in her lab, and enthusiastically supported my application to graduate school. The Hooper Foundation was a terrific place to work, both because of the superlative views of the Marin Headlands, the Golden Gate Bridge and the ocean from within a few steps of my lab bench, and because of the wonderful friends I made there. Roland Bainton, Owen Hughes, Lidia Arcizewska, Stephanie Wright and Ken Kubo were always valuable sources of scientific information, and great friends in and out of lab.

I thoroughly enjoyed my rotations in the Grosschedl, Walter, and Guthrie labs and learned valuable scientific lessons during each. However, the bulk of my time at UCSF and in graduate school has been in the laboratory of Elizabeth Blackburn, and for this I am enormously grateful. Liz has been an outstanding scientific mentor. She has allowed me a great deal of freedom in my work, and has always been available to discuss interesting data and ideas. As the number of people in the lab working on my particular project waned Liz continued to be interested and enthusiastic about my progess. I particularly value Liz's excitement about science, the creativity of her thought, and her even temperment.

One of the most wonderful aspects of Liz's lab has been the camaraderie and companionship I have found among the many people I have worked with over the years. Jeff Kapler provided me with an excellent example of how to be a careful scientist, and I gained a great deal from our collaboration in the lab. I enjoyed our ski trips and our shared interests in boating. Phil Cohen got me to begin footprinting the mutants in chromosome maintenance Jeff had turned over to me. That became a large portion of my thesis work and generated some very interesting results. Phil always took a paternal interest in my progress and seemed to want to be sure I wasn't in graduate school as long as he was. I wasn't, I was in graduate school longer!

In more recent years the friendship I have shared with Cathy Strahl, Anat Krauskopf and Karen Kirk has been a wonderfully

enriching and sustaining part of my life. It is unfortunately no exaggeration to say that we have shared great joys and great tragedies. I can't imagine the last several years without their concern and companionship. I am already anticipating the pain of losing their daily presence in my life. John Prescott has been a great benchmate and a great friend, so too have Anamitra Bhattacharrya and He Wang, and the other members past and present of the lab.

I would like to thank the members of my thesis committee. I have incredible admiration for Bruce Alberts, who taught me to focus on the big picture (though I may not have learned the lesson). He has an incredible ability to see the essential point of an issue and to refocus discussion on it, he is remarkable in his capacity to listen, and he never appears arrogant, despite his fantastic intellect. I owe many thanks to Don Ganem for encouraging my interest in science when I was a first and second year medical student, and continuing to support me when I decided to get really serious about it. Joachim Li took Bruce Alberts' place on my thesis committee and contributed his vast knowledge of DNA replication to my thesis committee meetings, and helped me expand my knowledge by allowing me to attend his lab's journal club. He was a very committed thesis committee member and my work and my thesis benefited greatly from his input.

There are not words to thank my partner Annie, without whom my life would be not nearly as much fun as she makes it, and without

whom producing this thesis would have been orders of magnitude more miserable than it was. She has stepped in at the darkest hour and led me forward again and again. This thesis would not have been submitted on time had it not been for the many hours of work she put in helping me pull it together. Her artistic abilities resulted in the final figure of Chapter 2, and her perfectionism resulted in the beautiful thesis you see before you. During the time in which I spent long hours in the lab completing experiments she was understanding and patient. When I did nothing but work on my thesis in the month before its submission she did my laundry, bought my groceries, and kept me sane. One of my main joys in completing this work is that it will allow me to spend more time with her, in more enjoyable occupations. She is been wonderful partner and has outdone herself in her contribution to my thesis. I can never repay her but I intend to try.

Structure and Function at a Replication Origin Region Within the *Tetrahymena thermophila* rDNA Macronuclear Chromosome

by
Renata Constance Gallagher

ABSTRACT

The 21 kilobase *Tetrahymena thermophila* macronuclear rDNA minichromosome is replicated from an origin of bidirectional replication located in the center-proximal 1.9 kilobase 5' non-transcribed spacer (NTS) of the rRNA genes. This is one of the best mapped eukaryotic cellular replication origins. I have taken advantage of a number of the unusual features of this chromosome to more fully characterize its structure and function.

I used DNase I, DMS, and $KMnO_4$ to modify DNA in chromatin in nuclei of cells, and determined the sites of modification precisely by primer extending modified DNA. This allowed me to demonstrate the presence of seven highly positioned nucleosomes which occupy two thirds of the 5' NTS. I identified identical, unusual footprints on two of these nucleosomes. Each is adjacent to a non-nucleosomal region which is a candidate site for the binding of origin recognition factors. This is the first description of an unusual footprint at an origin proximal nucleosome. I suggest that the unusual footprint is due to a distinctive structure that reflects a specialized role of the nucleosome in origin function, perhaps DNA unwinding. Thus, the

Tetrahymena rDNA origin may have a structure similar to that of ARS1 of *Saccharomyces cerevisiae*. There is a highly positioned nucleosome at the C region of ARS1; this region is required for ARS function in the absence of the B region of the ARS.

I describe the identification of two new mutants in rDNA minichromosome maintenance, one in the promoter region of rRNA genes, another in the origin region of the rDNA. I describe functional analyses of a maintenance mutant with a base change in the promoter region and demonstrate for the first time that an rDNA maintenance mutant (rmm) has a replication phenotype. The initiation of transcription was assayed in this mutant and there was no detectable phenotype. Genomic footprinting indicates that it has lost the wild-type footprint at the promoter at the site of the maintenance mutant base change, and at each of the two candidate origin regions 700 and 1100 base pairs away. Thus, the genetic evidence and footprinting evidence both indicate that the promoter region and the origin region are important for wild-type maintenance of the chromosome, and suggest that an interaction between the two is necessary for wild-type maintenance.

TABLE OF CONTENTS

CHAPTER 1 - Introduction

Tetrahymena thermophila as a Model System 2

Tetrahymena thermophila Histones and HMG-Like Proteins ..5

Causes and Occurrences of Nucleosome Positioning 10

Identification and High Resolution Mapping of Replication Origins 14

The Structure of Replication Origins and the Roles of Transcription Factors at Origins 22

Summary 24

CHAPTER 2 - Anatomy of A Replication Origin: The *Tetrahymena* Macronuclear rDNA Minichromosome Origin Region Contains Highly Positioned Nucleosomes, Origin Proximal Nucleosomes Display An Unusual Footprint

Abstract 26

Introduction 27

Results 35

Discussion 63

Materials and Methods 74

CHAPTER 3 - *Tetrahymena* rDNA Mutants Indicate that Interactions Between a Promoter Region and an Origin Region are Important for Wild-Type Chromosome Maintenance

Abstract 82

Introduction 83

Results 92

Discussion 125

Materials and Methods 140

REFERENCES 146

LIST OF FIGURES

Chapter 2

1a Structure of the *Tetrahymena* Macronuclear rDNA Minichromosome 29

1b Origin Region Protein-DNA Interactions 33

2a DNase I Footprinting of Nucleosome 2 on the Non-Coding Strand 38

2b DMS Footprinting of Nucleosome 2 on the Coding Strand 43

2c KMnO4 Footprinting of Nucleosome 2 on the Coding Strand 47

3a, 3b Low Resolution View of Nucleosomes 4 and 5 Footprinted with DNase I on the Coding Strand 50

4a, 4b High Resolution View of Nucleosome 4 on the Non-Coding Strand and of Nucleosome 5 on the Coding Strand Footprinted with DNase I 52

5 DMS Footprinting of Nucleosomes 5 on the Coding Strand 55

6a KMnO4 Footprinting of Nucleosome 4 on the Non-Coding Strand 58

6b, 6c KMnO4 Footprinting of Nucleosomes 4 and 5 on the Coding Strand 61

7 Sequence Alignment and Cleavage Sites 68

8 Sites of DMS and KMnO4 Reactivities on the Coding Strand of Nucleosomes 4 and 5 72

LIST OF FIGURES

Chapter 3

1 Structure of the Tetrahymena Macronuclear rDNA Minichromosome 85

2 rDNA Maintenance Mutants have Critical Base Changes Within the 5' NTS 89

3 An *In Vivo* Competition Assay Demonstrates the Maintenance Defect of C3-*rmm3* rDNA 92

4 2-D Gels Indicate Enhanced Accumulation of C3-*rmm3* rDNA Replicating Intermediates in 5' NTS 99

5a rDNA Chromosome Schematic: One Half of the Palindrome 105

5b *rmm3* Homozygotes Are Not Defective in the Initiation of Transcription 108

5c "Initiator Fragment" Transcripts Are Not Altered in C3-*rmm3* Cells 113

6a C3 Wild-Type rDNA Has a DMS Reactive A at Nucleotide 1132 in Domain 2 117

6b C3-*rmm3* rDNA Has No DMS Reactive A at Nucleotide 1132 in Domain 2 121

7a C3 Wild-Type DMS Promoter Footprint 124

7b C3-*rmm3* Lacks the Wild-Type DMS Promoter Footprint 127

8a C3 Wild-Type DNase I Promoter Footprint 129

8b C3-*rmm8* Has a Nearly Wild-Type DNase I Promoter Footprint 131

9 Model of Interaction Important for Wild-Type rDNA Maintenance 137

CHAPTER 1:

Introduction

Tetrahymena thermophila as a Model System

There are 7300 species of ciliated protozoa in 5 orders and 23 genera. All ciliates are single-celled organisms with two types of nuclei, the germ-line micronucleus and the somatic macronucleus. The macronucleus develops from a copy of the new micronucleus during mating in a process that involves chromosome breakage, DNA deletion, DNA rearrangement and gene amplification. The details of these processes differ between different orders. In the hypotrichous cilitates, for example, DNA deletion during macronuclear development is quite extensive, eliminating almost all but unique sequence DNA, and each gene is on an individual macronuclear chromosome. *Tetrahymena thermophila* is in the order of holotrichous ciliates. DNA deletion is relatively modest in these ciliates, so that it was not initially detected. About 10 per cent of micronuclear sequences are eliminated during macronuclear development. Thus, the number of macronuclear chromosomes differs greatly in different ciliates. In addition, the ploidy of individual chromosomes is very different in different ciliates. In *Tetrahymena thermophila* the ploidy of most of the two hundred macronuclear chromosomes is about fifty, with the exception of the rDNA minichromosome which is present at ten thousand copies per cell. In other ciliates the rDNA can be present at one hundred thousand copies per cell, and other chromosomes present at ten thousand copies per cell (Elliot, 1973; Gall, 1986; Prescott, 1994).

The unusual biology of ciliates, particularly with respect to nuclear differentiation and development, make them an interesting model system in which to study DNA rearrangement, DNA amplification, DNA deletion, and telomere addition, and other processes. However, ciliates diverged before the appearance of fungi (Prescott, 1994), and many highly conserved cellular proteins are less conserved between ciliates and other species than they are between organisms we consider to be quite divergent, yeast and humans for example. This will be discussed below with respect to histones and High Mobility Group (HMG)-like proteins, and is also true for TATA-binding protein (Bustin et al., 1990; Gorovsky, 1986; Stargell and Gorovsky, 1994). Thus, is it can be difficult to extrapolate from results in other organisms to *Tetrahymena*, or from *Tetrahymena* to other organisms. Comparison of rRNA sequences reveals that evolutionary divergence between ciliates can also be quite great, such that some ciliates are less related to each other than they are to other organisms (Prescott, 1994).

Tetrahymena thermophila is one of the most well studied ciliates and has been used for studies of cell morphology, DNA metabolism, gene expression, telomere addition, DNA rearrangement, and chromatin structure (Gall, 1986). I shall describe some of the unusual biological features of *Tetrahymena,* and some important experimental considerations. There are seven different mating types, under conditions of starvation cells of different mating types can mate. Macronuclear development is complete about twenty-four hours after pairing of starved cells (Gall, 1986). *Tetrahymena* are

grown in liquid culture, at 30°C, with aeration, in defined or rich media. Vegetative cells have a generation time of 3 hours under normal conditions. *Tetrahymena* has a variant genetic code, there are only two stop codons for translation. UAG codes for glutamine in *Tetrahymena,* making expression of *Tetrahymena* proteins in other organisms difficult, and expression of proteins of other organisms in *Tetrahymena* also a challenge (Prescott, 1994). In addition, transformation of *Tetrahymena* has not been well-developed, there are few vectors and only electroporation into the macronucleus of mating cells is efficient. Transformation of vegetative cells has been done by microinjection and is very labor-intensive (Gaertig and Gorovsky, 1995). Germ-line transformation of the micronucleus has been reported but has not yet been published (Peter Bruns, personal communication).

This brief overview of the biology and experimental considerations of *Tetrahymena thermophila* make clear that these ciliates are an important model system for a variety of investigations, and that there are great advantages and some disadvantages to *Tetrahymena* as an experimental system.

Tetrahymena thermophila Histones and HMG-like Proteins

As mentioned above, ciliates are quite divergent from other organisms. This is best illustrated for the histones, which are often cited as among the most highly conserved proteins of eukaryotes. The histones of ciliates have been characterized most extensively in *Tetrahymena* species (*thermophila and pyriformis*). Among the histones, H4 is the most highly conserved. It is at least 90 per cent conserved between plants, animals and fungi. In contrast, *Tetrahymena* histone H4 is only 80 per cent conserved with those of these groups, and could not replace calf histone H4 in reconstitution experiments. It is the same length, 102 amino acids, as the H4 of other organisms, but has an unblocked alanine, instead of acetylserine at its N-terminus. There are two genes for H4 in *Tetrahymena thermophila*. These encode the identical protein, the genes differ in their expression. One is constitutive and one is cell cycle regulated (Gorovsky, 1986; Yu et al., 1987).

There are three genes for histone H3 in *Tetrahymena*. Two of these genes encode the same protein, the third encodes the variant hv2, which resembles the minor mammalian H3.3. The hv2 protein is minor in *Tetrahymena*, and is synthesized and deposited in macronuclei of growing and non growing cells (Allis et al., 1980). The major H3 is also more divergent between *Tetrahymena* and other organisms than it is among those organisms, it is 20 per cent divergent from the histone H3 of *S. cerevisiae* (Gorovsky, 1986).

There are three H2A genes. Two differ from each other slightly in sequence and both are quite divergent from the H2As of other organisms, they are 25 per cent divergent from *S.cerevisiae* H2A. The third encodes a variant, hv1, that is remarkably similar to H2A variants found in chicken, mammals, sea urchins and flies. These variants are more similar to each other than they are to the H2As of other organisms, indicating that their specialized function evolved early in evolution and has been preserved. In *Tetrahymena* this variant is found specifically in the transcriptionally active macronucleus, suggesting it plays a role in the assembly of transcriptionally competent chromatin. Further evidence in support of this came from the observation that hv1 is present in micronuclei early in conjugation, coincident with the brief period in which the micronucleus is transcribed. It disappears from micronuclei as the new macronucleus becomes competent for transcription (Stargell et al., 1993). There is only one H2B gene, this is also the most evolutionary divergent of the known H2B histones, it is 36 per cent divergent from *S.cerevisiae* H2B. However, this histone can substitute for calf H2B in nucleosome reconstitution (Gorovsky, 1986).

Tetrahymena macronuclear H1 is unusual. It lacks the central hydrophobic domain conserved in all other H1s, and it has an unusually large number of positively charged amino acids. There is a single gene for this histone. Like the H1 of other organisms, it does dissociate from chromatin at lower salt concentrations than the other histones, and it is located in the linker regions of chromatin.

Micronuclei lack histone H1. They have a set of polypeptides, processed from a larger gene, that are associated with the linker regions of chromatin and lack typical H1 properties. Thus, the chromatin of transcriptionally active macronuclei and transcriptionally inactive micronuclei differ with respect to the presence of at least two major histones. This is believed to be related, at least in part, to their differing transcriptional states (Gorovsky, 1986).

High mobility group (HMG) proteins are the largest class of non-histone chromatin associated proteins. They are believed to play a role in chromatin structure, gene expression, DNA replication, and recombination. They are identified by their physical properties: solubility in 2 to 5 per cent perchloric acid, extractability from chromatin at 0.35 M NaCl; a molecular weight of less than 30 kilodaltons; and a high content of charged amino acids (Bustin et al., 1990). In higher eukaryotes these proteins are grouped into three classes. These are the HMG 1/2, HMG 14/17, and HMG I/Y groups. The precise role of these groups has not been defined. HMG 1/2 proteins have been shown to bend DNA (Paull et al., 1993); HMG 14/17 proteins have a higher affinity for nucleosome cores than naked DNA and bind the nucleosome at two sites, they are believed to play a role in transcription initiation and/or elongation (Alfonso et al., 1994); HMG I/Y proteins are found specifically in rapidly dividing cells and may be associated with H1 depleted chromatin (Zhao et al., 1993). Four proteins with HMG-like properties have been isolated from *Tetrahymena*. Two of these, HMG B and HMG C, have been well.

characterized. They have an HMG box and appear to be most closely related to the vertebrate HMG 1/2 family by sequence, but to the HMG 14/17 family in their physical properties (Schulman et al., 1991). They are present in both micronuclei and macronuclei, but are much more abundant in macronuclei. HMG B is induced early in conjugation and may play a specific role in one or more of the processes that occur in this period, replication, recombination, repair, transcription (Wang and Allis, 1993).

In summary, the nuclear dimorphism of ciliates makes the study of chromatin structure and protein composition in these organisms particularly interesting and likely to reveal insights into the relationships between chromosome structure and function, and protein composition and function. The histone proteins of *Tetrahymena* are the best studied of any ciliate, and recently more direct attempts to determine their function in micronuclei and macronuclei have come from gene knockout experiments (Shen et al., 1995). While the histones and HMG-like proteins of *Tetrahymena* are quite divergent from those of other organisms, insights gained in this system are likely to shed light on the role of these components in other organisms, either as a result of striking similarity, as in the case of the histone H2A variant hv1, or difference, as in the case of the unusual macronuclear H1 and very unusual micronuclear linker-associated polypeptides.

In chapter two of my work I have defined the chromatin structure of a macronuclear replication origin region at high resolution. I have

shown that in spite of the great divergence of *Tetrahymena* histones, the DNase I footprints of *Tetrahymena* macronuclear nucleosomes are indistinguishable from those of other organisms, with a regular 10 base pair periodicity of DNase I cleavage. In addition, I have found that nucleosomes at the origin region are very highly positioned. Interestingly, I have found unusual footprints at two origin proximal nucleosomes and I suggest that this may reflect an altered nucleosomal structure that facilitates DNA replication.

In subsequent work it will be of interest to determine whether this region is active as an origin in micronuclei and/or amplifying cells, to define the chromatin structure of this region in micronuclei and in amplifying cells, to identify the specific modification or protein(s) responsible for the unusual footprint at origin proximal nucleosomes, and to determine the role of this modification in DNA replication and nucleosome structure. The unusual footprint includes DMS reactive A residues, which are not found in purely nucleosomal DNA (McGhee and Felsenfeld, 1979). HMG proteins bind in the minor groove of DNA and interact with the N-3 positions of adenines (van de Wetering and Clevers, 1992). Thus, candidates for the footprint I have identified include the HMG-like proteins of *Tetrahymena,* HMG B and HMG C.

Causes and Occurrence of Nucleosome Positioning

Nucleosome positioning has been observed at a variety of locations. It can be caused by several different mechanisms (Simpson, 1992). A specific DNA sequence may adopt or energetically favor a structure to which nucleosomes bind with a specific conformation. The 5S rRNA gene of several organisms has been shown to have such a positioning sequence (Buttinelli et al., 1993; Simpson and Stafford, 1983). Alternatively, nucleosomes may be highly positioned in a region because a protein boundary precludes them from moving in one direction. This may explain the positioned nucleosomes extending in from the telomere in *Tetrahymena* (Budarf and Blackburn, 1986). A highly ordered array of nucleosomes also extends outward in both directions from the centromere of chromosome III in S. cerevisiae, but this has been shown to be independent of the binding of centromere proteins and is more likely due to a positioning sequence in the DNA (Bloom and Carbon, 1982). A particular DNA sequence might form a structure that specifically excludes nucleosomes, this would be a boundary established by DNA. It has also been suggested that proximity to a replication origin might result in positioned nucleosomes, in this case the origin would be the boundary establishing nucleosome positioning (Simpson, 1992). A specific protein interaction may position a nucleosome as a result of or in order to facilitate a process in the cell. For example, nucleosome positioning at the albumin enhancer is found in cells in which the albumin gene is expressed, but nucleosomes are found randomly at this location in cells in which the gene is inactive. Specific factors are bound to the nucleosomes

in cells expressing albumin. Nucleosome positioning may arise as a consequence of the binding of these factors, or may allow the binding of these factors (McPherson et al., 1993).

There are two types of nucleosome positioning. The first is translational positioning. This refers to the 5' and 3' locations of the DNA wrapped around the nucleosome, either 145 base pairs in nucleosomal cores, or 166 base pairs in nucleosomes plus H1 (chromatosomes). Rotational positioning refers to the number of sites on the nucleosome which contact DNA. For example, DNase I cleaves nucleosomal DNA at 10 base pair intervals, where the minor groove is exposed on the surface of the nucleosome. In a population of molecules bound by nucleosomes that were not rotationally positioned, DNase I would cleave DNA at every available site. In a population of molecules in which there is exact rotational positioning DNase I will cleave DNA at precise 10 base pair intervals. Precise rotational positioning can occur in a population of nucleosomes in which there are multiple translational positions, each rotated by ten base pairs from the other. This was documented in vitro on nucleosomal cores positioned on the 5S rRNA gene in the absence of linker histones (Pennings et al., 1991). A population of mucleosomes that are all at exactly the same translational postion will also share a unique rotational position. However, in a population of nucleosomes that are precisely rotationally positioned, individual nucleosomes may have different translational positions, these will be staggered by a multiple of 10 bps from the others. Thus, translational and rotational positioning of

nucleosomes are not independent, but they are interrelated. Linker histones are thought to contribute to nucleosome positioning, perhaps by locking rotationally positioned nucleosomes into one of several alternate translational positions (Meersseman et al., 1991).

Positioned nucleosomes have been found in the vicinity of the best described viral and cellular replication origins, SV-40 and ARS1, respectively (Jakobovits et al., 1980; Thoma et al., 1984). Similarly, the *Tetrahymena* rDNA origin region contains positioned nucleosomes flanking non-nucleosomal regions that are good candidates for the sites at which origin recognition factors might bind. The *Tetrahymena* origin region is located just off the center of the palindromic 21 kilobase (kb) minichromosome. This region contains the 1.9 kb 5' non-transcribed spacer (NTS) of the Pol I rRNA genes (17S, 5.8S, 26S), which are divergently transcribed from the center of the molecule. The origin region of the macronuclear rDNA minichromosome in *Tetrahymena* contains seven highly positioned nucleosomes, a DNase I sensitive region adjacent to the 5' end of the divergently transcribed pair of rRNA genes (located at the center-distal portion of the 5' NTS), and two non-nucleosomal regions (these are homologous, each is within one repeat of a tandemly repeated sequence) which are likely to be the sites of origin recognition factor binding (Palen and Cech, 1984) (this thesis). I have identified unusual DMS and KMnO4 reactivies within the nucleosomes adjacent to the non-nucleosomal regions. These unusual reactivities may be indicative of an altered structure at these nucleosomes. I suggest that the fact that the nucleosomal

regions have been retained in the tandem duplication, and the fact that they have unusual footprints, indicates that, together with the adjacent non-nucleosomal regions, they are part of structural and functional unit. I further suggest that the unusual reactivities may reflect an unusual structure that facilitates the function of this unit, replication. If these nucleosomal regions are involved in facilitating origin function, it could be through an altered structure which facitlitates origin unwinding. Alternatively, it is possible that these regions establish nucleosome positioning, keeping the origin region clear of nucleosomes.

The chromatin structure of the *Tetrahymena* origin region is very similar to that of ARS1 of the TRP1ARS1 plasmid. On this 1.4 kb episome which contains the TRP1 gene, the TRP1 transcription unit is flanked by DNase I hypersensitive sites at the 5' and 3' ends. Within the 3' site are the A and B regions of ARS1, originally defined by Celniker et. al. (Celniker et al., 1984). There are seven positioned nucleosomes on the plasmid. The C region of ARS1 corresponds to one of these (Simpson, 1990). C is required for ARS function in the absence of the B region located on the other side of the A consensus ARS element. It is an interesting possibility that the nucleosomes at which I have identified unusual footprints may be analogous to the C region of ARS1. The role of the C region is not clear. As suggested above for the *Tetrahymena* origin region, it could be involved in facilitating unwinding, or in establishing nucleosome positioning. It is also possible that nucleosomal positioning could be due to origin recognition factors binding and excluding

nucleosomes from the origin region. However, Simpson has shown that deletions within the TRP1ARS1 plasmid that position the ARS consensus sequence within the positioned nucleosome in the C region reduce ARS function (Simpson, 1990). This suggests that, at least for the TRP1ARS1 plasmid, nucleosome positioning is not secondary to origin recognition complex (ORC) binding. Further work will establish the role of the origin proximal nucleosomes in the *Tetrahymena* rDNA origin and the cause of their altered footprint, and the specific role of the C region in ARS1. These may be the same or different. It will be interesting to determine whether the C region nucleosome of ARS1 has unusual DMS and KMnO4 footprints similar to those I have identified in the *Tetrahymena* rDNA origin region.

Identification and High Resolution Mapping of Replication Origins

Crucial to cell viability and chromosome maintenance is the replication and transmission of genetic information. In chapter three of my thesis I describe the identification of rDNA macronuclear minichromosome maintenance mutants in *Tetrahymena*, and the functional analysis of one of these. Defective maintenance can be a function of defective replication, segregation, or both. These mutants display a maintenance phenotype in a nucleus which divides amitotically and in which faithful segregation and transmission of genetic information do not occur (Larson et al., 1991). I provide the first evidence that one of these mutants has a

replication phenotype, and suggest that the effects of the maintenance mutants are due to an effect on DNA replication.

The base changes associated with maintenance defects in these mutants are in the 5' non-transcribed spacer of the chromosome (See Figure 2, Chapter 3) (Larson et al., 1986; Yaeger et al., 1989) (this thesis). This is the location of the replication origin identified physically by electron microscopy (Cech and Brehm, 1981), and 2-dimensional gels (Kapler and Blackburn, unpublished work; this thesis), and functionally by transformation (See Figure 1, Chapter 2) (Gaertig and Gorovsky, 1995) (Meng-Chao Yao, personal communication). I describe the chromatin structure of the origin in chapter two of my thesis. This chromatin structure reveals that there are two non-nucleosomal domains that are good candidates for the sites at which origin recognition factors could bind. Here I will describe our attempts to identify the origin of replication in the *Tetrahymena* rDNA, and the structure of origins of replication in the organism in which they have been defined most comprehensively, *S. cerevisiae*.

Origins can be identified physically by methods that detect replicating molecules, such as 2-dimensional gels or electron microscopy, or origins can be defined functionally by assays which determine the sequences that allow stable maintenance of an episome. These may not identify the same regions, as in the Drosophila chorion genes in which the sequences responsible for DNA amplification are not the sites which direct this amplification.

However, for most of the origins in which both have been described, the physical and functional origins localize to the same region (DePamphilis, 1993). This also appears to be the case for the replication origin of the *Tetrahymena* rDNA minichromosome. Meng-Chao Yao has reported that a 900 base pair region which includes the two tandem repeats within the 5' NTS and approximately 100 base pairs of flanking sequence will allow an episome transformed into mating cells to be maintained (Meng-Chao Yao, personal communication). However, assaying maintenance of plasmids transformed into mating cells does not allow the distinction between sequences required for amplification of the molecule to ten thousand copies during macronuclear development, and sequences required for vegetative maintenance of this episome. I attempted to transform vegetative cells by electroporation in order to directly address the location of the functional vegetative origin. I employed a transient transformation assay which would distinguish replicated molecules from input DNA by their methylation state. It was necessary to use PCR to amplify DNA from cells electroporated and grown for twenty-four hours in order to identify plasmid DNA. Initial results were promising, but the definitive experiment demonstrated that there was no difference in my assay between a plasmid lacking an rDNA origin and a plasmid containing an rDNA origin. Therefore, the minimal functional rDNA origin identified to date does not distinguish between the amplification and vegetative origins.

Using 2-dimensional neutral-neutral gels Jeff Kapler attempted to determine whether the physical amplification and vegetative origins are coincident (Kapler and Blackburn, unpublished results). The early amplification origin detected at twelve hours after mating is initiated is located outside the 5' NTS, and there appear to be multiple initiations on the same molecule. By fourteen hours post-pairing there is evidence of firing of the origin within the 5' NTS (presumably the same as the vegetative origin). By eighteen hours there is a 50:50 mixture of both patterns. At twenty-four hours the pattern is equivalent to that see in vegetative cells (Jeff Kapler, unpublished results). Thus, it appears that the early amplification origin is different from the vegetative origin, but it is not yet clear whether that origin is as defined as the vegetative origin appears to be.

There are many issues to be addressed regarding the functions of the two origins. There may well be a role for chromatin structure, or its absence, at early timepoints of macronuclear development. In addition, transcription of the rRNA genes may play a role in determining origin location and function. The act of transcription may restrict origin firing from the transcription unit in vegetative and late amplifying cells directly, or the active promoter may bind proteins that have a replication enhancer effect (this thesis). The organization of chromatin structure of the rDNA in the micronucleus or during macronuclear development is not known. rRNA transcription begins at around 16 hours after mating is initiated. Similar developmental issues have been addressed for the rDNA

origins of Xenopus (Hyrien et al., 1995). These authors observe a marked decrease in origin firing from the transcription units of rRNA genes as these genes become competent for transcription. However, they conclude that the decrease in origin firing is due to chromatin remodeling, not transcription itself, because the reduction in firing is greater than they expect from the amount of transcription they believe is taking place. One difficulty in determining structure/function relationships in the rDNA is that it is difficult to determine definitively that the molecules one is assaying for structure are all active, or are the same population of molecules one is assaying for function.

The nine-hundred base pair region to which the physical and functional origins of the *Tetrahymena* rDNA minichromosome have been mapped include two nucleosomal regions. It is most likely that origin recognition factors bind in the non-nucleosomal regions. Together these comprise under six hundred base pairs of sequence. Since each domain is within one copy of a tandemly repeated sequence, the region likely to bind an origin recognition factor is between two hundred and fifty and three hundred base pairs. This is a relatively small region, but still large compared to the 11 to 15 base pair sequence required for binding the origin recognition complex (ORC) of *S. cerevisiae* (Diffley and Stillman, 1990). The *cerevisiae* ARS consensus sequence was initially determined by comparison of those sequences which would allow maintenance of a yeast episome. The sequence itself is not sufficient for plasmid maintenance, but it is necessary for plasmid maintenance. Similar

functional analysis of the vegetative origin is difficult in *Tetrahymena* because there is no efficient way to transform vegetative cells. As an alternative to precisely defining the functional origin attempted to precisely define the physical origin. The physical method that allowed determination of the SV-40 and polyoma viral replication origins to the nucleotide should, in theory, be applicable to the *Tetrahymena* rDNA origin. This assay is the Okazaki fragment strand switch assay analyzed at high resolution. Okazaki fragments are the short DNA fragments produced by lagging strand synthesis which are ligated together to complete DNA replication of that strand. The location of the OBR determines which sequences are replicated by leading strand syntheseis and which by lagging strand synthesis. The lagging strands on opposite sides of the OBR are on opposite strands of the DNA. The strand switch assay utilizes hybridization of Okazaki fragments to single strand origin region DNA probes. By determining where Okazaki fragments stop hybridizing to one DNA strand and begin hybridizing to the other it is possible to localize the OBR. Okazaki fragments were isolated and hybridized to single stranded SV-40 origin region DNA. The hybrid was restriction digested and sized on sequencing gels, allowing precise determination of the location of the 5' end of the Okazaki fragment (Hay and DePamphilis, 1982). Since the general region of the physical origin of the *Tetrahymena* rDNA is known, and probes are available for the entire region, it should be possible to do this assay. However, the experimentally observed strand bias in the Okazaki strand switch assay is only four-fold at best. Since the *Tetrahymena* rDNA minichromosome is a palindrome replicated

bidirectionally, one half of the palindrome will always be passively replicated, this will practically obliterate the strand switch signal (Bill Burhans, personal communication). Another technique that could allow a more precise determination of the rDNA origin are neutral-alkaline two-dimensional gels. In this technique replicating molecules are run in the first dimension for size, and in the second dimension in an alkaline buffer which denatures nascent strands. Different probes are hybridized sequentially to the nascent strands. The probe which detects the smallest strand is closest to the origin. This technique can be used to determine the direction of replication fork movement, and was used to localize the bovine papilloma virus (BPV) origin to within 100 base pairs (Nawotka and Huberman, 1988; Yang and Botchan, 1990). To date, it has not been possible to detect small nascent strands from the *Tetrahymena* rDNA origin with this technique. Instead these gels identify long molecules which have been initiated on the other side of the palindrome, and are stalled at replication pause sites first identified by neutral-neutral 2-dimensional gels. This has allowed localization of the pause sites, but not a determination of the origin (Jeff Kapler, personal communication).

The structure of the SV-40 and ARS1 origins suggest that origins are likely to be located in nucleosome free regions. Thus, footprinting has localized the *Tetrahymena* rDNA origin region most precisely, by clearly delineating those non-nucleosomal regions most likely to bind origin factors (Palen and Cech, 1984) (this work). It does remain a possibility that the nucleosomes of the origin

region are cleared during the cell cycle from a population of molecules that is not detectable in my work, and that the nucleosomal regions bind origin recognition factors. However, it would then be necessary to explain the function of the nucleosome free region of the origin region. Thus, by analogy with other well-studied origins it is most likely that the nucleosome free regions bind origin recognition factors. The only eukaryotic cellular origin recognition factor identified to date is the origin recognition complex (ORC) of *S.cerevisiae* which binds the ARS consensus A element (Diffley et al., 1995). The role of this complex in DNA replication has not yet been determined. *In vivo* footprinting reveals that ORC is bound throughout the cell cycle, and the footprint of the complex changes during the cell cycle. ORC has a characteristic DNase I footprint on DNA. No other reagents have been reported as having been used to footprint ORC. No DNase I footprint analogous to that of *S. cerevisiae* ORC was found in the non-nucleosomal regions of the *Tetrahymena* rDNA. In fact, naked DNA and chromatin had exactly the same pattern and intensity of cleavages in both non-nucleosomal regions, that is, there was no DNase I footprint at all. However, in chapter three of my thesis I describe a DMS reactive A residue in the non-nucleosomal region that was not present in a chromosome maintenance mutant. This nucleotide is located 61 nucleotides upstream of a sequence identified through other mutants as playing a role in chromosome maintenance (Larson et al., 1986; Yaeger et al., 1989) (this thesis). Its presence in the non-nucleosomal region and its specific absence in a maintenance mutant suggests that it indicates the binding of a core or auxiliary

replication factor, or a DNA structure important for replication function.

The Structure of Replication Origins and the Roles of Transcription Factors at Origins

In chapter three I describe that footprinting and the location of maintenance mutants also suggest that an interaction between the origin region and the promoter region of the rDNA is important for wild-type chromosome maintenance. A role for transcription factors in DNA replication has been identified in numerous origins of replication. In the most thoroughly analyzed cellular origin, ARS1, ARS activity can be reconstituted from synthetic matches to the A region containing the ARS consensus sequence and three elements within the B region. Any two of the B elements are required for function, and at least two different elements are required. The B3 element binds the transcription factor ABF1. Other transcription factor binding sites can be substituted for ABF1 if their corresponding factor is present in the cell (Marahrens and Stillman, 1992). ABF1 can function in a distance and orientation dependent manner to enhance replication (Walker et al., 1990), however, it is not yet known how ABF1 functions. In other instances the role of transcription, or transcription factors has been better defined. In SV-40 the role of transcription factors appears to be to help keep the origin region free of nucleosomes (Cheng and Kelly, 1989). In adenovirus, in which DNA replication is primed by protein, two transcription factors enhance replication, NFI or CTF, and NFIII or oct-1. NFI facilitates initiation complex assembly, NFIII appears to

alter DNA structure at the origin. In each case the transcriptional activation domain is not required for origin enhancer function (Heintz, 1992). In other examples the transcriptional activation domain of a transcription factor is required for replication enhancer function. Different transcriptional activation domains have different roles in transcription and probably have different roles in replicaiton. *In vitro* results indicate that one possible function of a transcriptional activator domain is recruitment to the origin region of the eukaryotic single-stranded binding factor RPA (He et al., 1993; Li and Botchan, 1993). In BPV the E2 transcription factor may provide a bridge between the E1 origin recognition factor and replication enzymes (Heintz, 1992).

Thus, the probable roles of transcription factors in DNA replication include excluding nucleosomes, altering DNA structure, and interacting with essential and auxiliary replication factors. The role of the specific factor involved must be determined for each origin. One relationship between transcription and replication in eukaryotes is that transcribed regions are replicated early in S phase. This could be due to an open chromatin conformation of transcribed regions, or a more specific interaction. In the rDNA of *S. cerevisiae* Miller and Kowalski have defined a minimal origin of 107 base pairs, which includes an ARS element that is not a good match to the consensus and a broad 3' region. There is no ABF1 binding site in the minimal origin (Miller and Kowalski, 1993). It is possible, and is suggested by our results in *Tetrahymena*, that the Pol I promoter of the rRNA genes itself provides a replication enhancer function,

and links transcription and replication in the rDNA. The *S. cerevisiae* rDNA ARS is weak, and only about one in five rDNA replication origins fires. Perhaps those units in which an origin fires are also those which are actively transcribing their rDNA. It is important to note that unlike *Tetrahymena* the 5' NTS of the *S. cerevisiae* rDNA includes the Pol III transcribed 5S rRNA gene, and this element may also contribute to rDNA replication.

SUMMARY

In this introduction I have described some of the background that has provided the foundation for my thesis work. I have provided a description of the biology of *Tetrahymena*, and its similarities and differences with other organisms. I have described important features of chromatin structure and DNA replication. The following chapters detail my investigations describing the chromatin structure at a replication origin in *Tetrahymena*, and my analysis of a chromosome maintenance mutant which has a replication phenotype and a base change at a promoter region adjacent to the origin region. Studies in *Tetrahymena* have elucidated fundamental aspects of biology. Here I have used it to provide an important example of structure and function at a replication origin.

CHAPTER 2:

Anatomy Of A Replication Origin: The *Tetrahymena* Macronuclear rDNA Minichromosome Origin Region Contains Highly Positioned Nucleosomes, Origin Proximal Nucleosomes Display An Unusual Footprint

Abstract

The 21 kilobase palindromic rDNA minichromosome of *Tetrahymena thermophila* is replicated from an origin within the center proximal non-transcribed spacer (NTS) of the rRNA genes. This 1.9 kilobase (kb) 5' NTS is located upstream of each of the two Pol I transcription units on the minichromosome. Genomic footprinting of the 5' NTS reveals that seven highly positioned nucleosomes occupy two-thirds of this origin region. There are two non-nucleosomal regions which are good candidates for the sites of origin recognition factor binding. Each is within a four hundred base pair sequence found in tandem direct repeat within the 5' NTS. These two repeats are highly homologous with respect to DNA sequence. We show that the two repeats are virtually identical in their footprinting properties, even in areas of decreased DNA sequence homology. Structurally, each repeat contains a highly positioned nucleosome at its 5' end, followed by roughly two-hundred and fifty base pairs of nucleosome free DNA. We demonstrate that the nucleosomes within these repeats have unusual, and identical, footprints. These consist of three DMS reactive A residues, and three strong KMnO4 sites which are not found in other nucleosomes, and do not correspond to the sites predicted from the structure of the nucleosome to be unstacked. We suggest that these origin proximal nucleosomes have a specialized structure due to a distinctive role in DNA replication, perhaps DNA unwinding, and that each approximately four hundred base pair repeat constitutes a structural and functional unit. The is the first report of the detailed structure of an origin proximal nucleosome. In

addition, we report the first genomic footprinting of nucleosomal DNA with KMnO4.

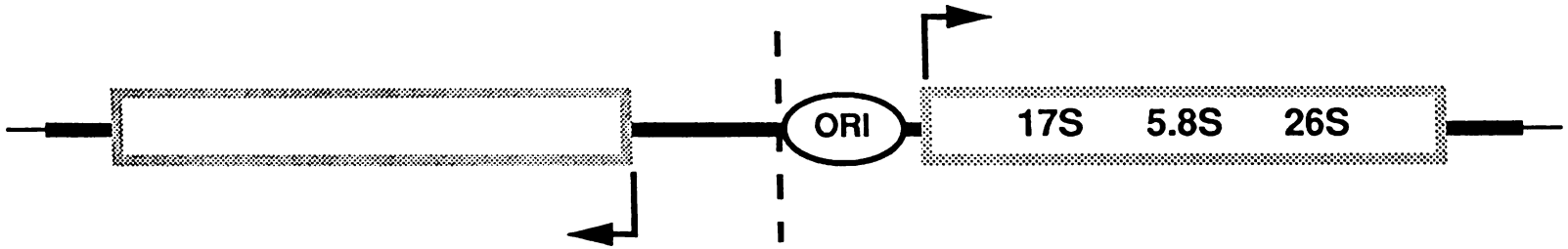
Introduction

The *Tetrahymena* macronuclear rDNA minichromosome contains one of the best localized eukaryotic, non-viral, DNA replication origins. This 21 kilobase (kb) palindromic minichromosome is replicated from an origin of bidirectional replication located near the middle of the palindrome, within the 1.9 kb 5' non-transcribed spacer (5' NTS) of the rRNA genes (Figure 1a) (Cech and Brehm, 1981; Truett and Gall, 1977) (Kapler and Blackburn, unpublished results). It is the smallest of the *Tetrahymena* macronuclear chromosomes, each 10 kb half-palindrome is comprised of a single Pol I transcription unit with flanking 5' and 3' non-transcribed spacers (Engberg and Nielsen, 1990). It is also the most abundant of the macronuclear chromosomes. The rRNA genes of most other organisms are found as tandem repeats; however, ciliated protozoa amplify all macronuclear chromosomes, and selectively over-amplify the rDNA chromosome. *Tetrahymena* has two hundred macronuclear chromosomes, most of these are present at only about fifty copies per cell. The rDNA minichromosome is present at ten thousand copies per cell (Prescott, 1994). Replication of this chromosome is largely under cell cycle control (Larson et al., 1986; Truett and Gall, 1977). The small size and the abundance of this chromosome contribute to its value as a model system in which to study a well-defined, cellular, eukaryotic origin of replication. We have undertaken both structural and functional studies of this origin

Figure 1a: Structure of the *Tetrahymena* Macronuclear rDNA Minichromosome

This 21 kb molecule is a palindrome, the dashed line marks its center. In most cases the two halves of the palindrome are identical, except for a 28 base pair non-palindromic region at the center of the molecule. The important structural features of the molecule are illustrated for the right half of the palindrome, and are also found on the left half. The heavy black lines indicate the 5' non-transcribed spacers at the center of the molecule, and the 3' non-trancribed spacer at the ends of the molecule. The thin black line represents telomeric DNA. The bent arrow represents the start site of transcription of the rRNA genes. There is a single Pol I transcription unit per half palindrome, its extent is indicated by the open rectangle. The 35S transcript produced from this region is processed into the 17S, 5.8S and 26S rRNA genes. The order of these within the transcription unit, but not the precise location, is indicated within the rectangle on the right. Finally, each half of the palindrome contains an origin of DNA replication (ORI) within the 5' NTS, indicated by the open oval for the right half of the palindrome.

Structure of the Tetrahymena Macronuclear rDNA Minichromosome



region (Kapler and Blackburn, 1994; Kapler et al., 1994; Larson et al., 1986; Pan and Blackburn, 1995; Pan et al., 1995; Yaeger et al., 1989) (this thesis).

Sequences that constitute an origin of DNA replication can be identified functionally, through plasmid stability assays, or physically, using a variety of techniques. The two types of assays may not identify the same DNA sequences as the origin (DePamphilis, Annual Rev.). However, for the cellular replication origins in which the physical and functional origins have been best defined, the ARS elements of *S.cerevisiae,* the functional ARS (autonomously replicating sequence) element and the physical origin of bidirectional replication (OBR) are in the same region, within the level of resolution of the physical assay (Brewer and Fangman, 1987; Huberman et al., 1987). A protein complex which binds the ARS element *in vitro* and footprints it *in vivo* has been identified. The footprint of this complex changes with the cell cycle, but its precise role in DNA replication has not yet been defined (Diffley and Cocker, 1992; Diffley et al., 1995; Diffley et al., 1994; Hernandez, 1993).

In *Tetrahymena* the physical origin of bidirectional replication has been mapped by electron microscopy to 600 +/- 300 base pairs off the center of the palindrome (Cech and Brehm, 1981). This was done in the B strain allele, which has a maintenance disadvantage in an *in vivo* competition with the C3 strain rDNA allele. The experimental observation is that in micronuclear C3/B heterozygous strains in which both alleles of the rDNA (distinguished by

restriction fragment length polymorphisms) have been amplified in the developing macronucleus, the B rDNA allele is specifically lost from the macronucleus within 70 to 100 generations after cells reach maturity, even when they are under selective pressure (drug resistance) for the presence of this allele. This maintenance disadvantage could be due to a segregation defect, a replication defect or a combination of the two. Recombinant chromosomes demonstrate that the base changes responsbile for the maintenance defect are a 42 base pair deletion adjacent to the Type I repeat element of Repeat 2 of the replication origin region (See Figure 1b). This observation implicates origin function in the defect. One hypothesis is that both Repeats 1 and 2 function as origins in the C3 strain, but that only Repeat 1 functions in the B strain as a result of the deletion in Repeat 2. Neutral-neutral two-dimensional gels of both the C3 and B strains have failed to distinguish a difference between the physical origins of the two alleles and have failed to determine whether the OBR is within the first of the tandem repeats (Repeat 1- located from nucleotides (nts) 503 to 935, see Figure 1b) or the second of the tandem repeats (Repeat 2-located from nts 935 to 1337, see Figure 1b). However, these gels do localize the OBR to the 900 base pair region of the two tandem repeats, and reveal the presence of strong replication fork pauses within the 5' NTS that may make it difficult to map the physical origin more precisely (Kapler and Blackburn, unpublished results). Results of transformation experiments indicate that the functional origin also localizes to the region of the two tandem repeats, and includes

Figure 1b: Origin Region Protein-DNA Interactions

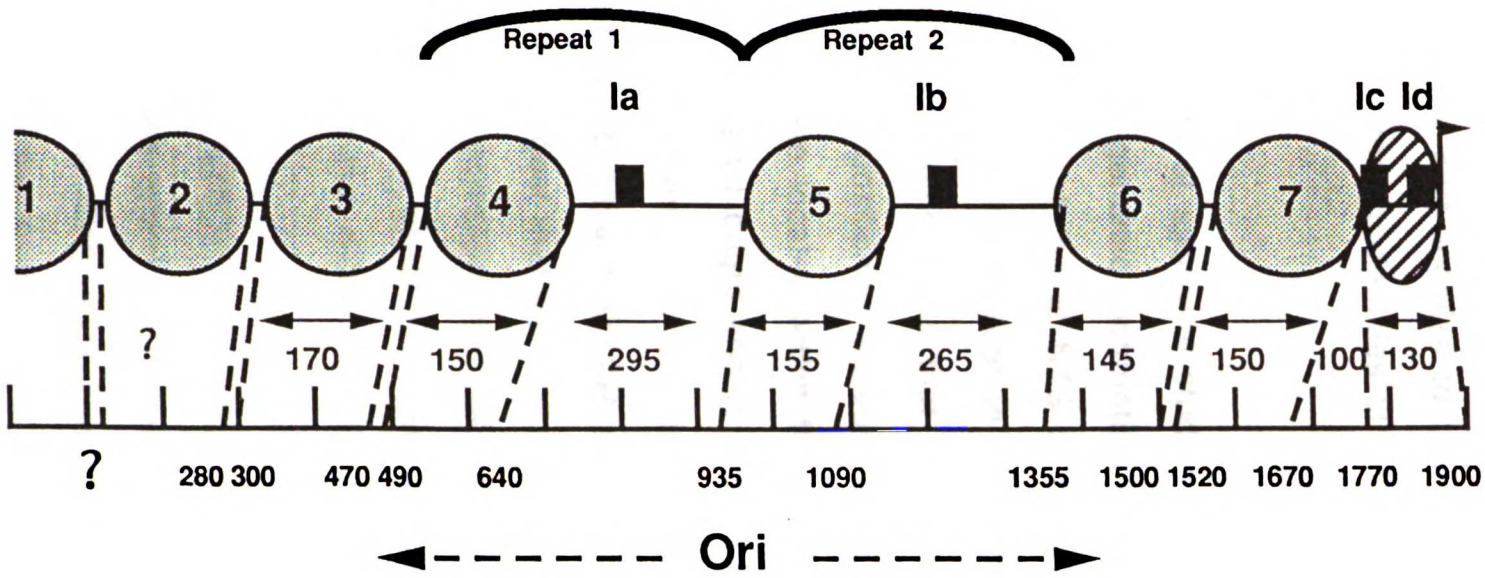
The rDNA minichromosome central 5' NTS is 1.9 kb and contains both the genetic and physical origins of the chromosome. A schematic of protein-DNA interactions is indicated above, below a scale bar marks the positions of these regions along the DNA. DNase I footprinting of nuclei reveals three classes of regions within the 5' NTS. There are seven highly positioned nucleosomes, these are indicated by the shaded circles. There are regions in which the cleavage pattern of naked DNA is equivalent to that of chromatin, these are indicated by a plain black line. There is a broad region of protections and enhancements near the start site of transcription, this is indicated by the hatched oval. The nucleosomal borders are approximate and are deduced from the presence of unprotected (linker) DNA between nucleosomes, and the sites and pattern of DNase I cleavages. Because DNase I does not cut at every nucleotide the accuracy of the nucleosomal border is limited to roughly +/- 10 base pairs.

Repeats 1 and 2 indicate the roughly 400 base pair domain found in tandem repeat in the 5' NTS.

Ia, b, c, and d represent 33 nucleotide sequence elements found in the 5' NTS of several different Tetrahymenid ciliates. They are conserved in location and sequence and are the sites of base changes in chromosome maintenanace mutants.

The start site of transcription is indicated by the flag it is at 1887.

Origin Region Protein-DNA Interactions



sequences just 3' of Repeat 2 (Meng-Chao Yao, personal communication). Thus, although the physical and functional origin(s) of the *Tetrahymena* macronuclear rDNA minichromosome have not been precisely delineated, they both localize to the region of the two tandem repeats. These span from nucleotides 503 to 1337, (.8 kb) of the 1.9 kb 5' NTS.

Here we report the results of structural studies of the entire origin region. We have used genomic footprinting to characterize protein-DNA interactions throughout the 5' NTS in order to identify regions likely to be important for the replication and maintenance of this chromosome. Our results demonstrate that two thirds of the 5' NTS is covered by seven highly positioned nucleosomes. Two of these, nucleosomes 4 and 5, are located within the Repeat 1/Repeat 2 region that contains the physical origin and is part of the functional origin. These two nucleosomes are located in highly homologous DNA sequences and have distinctive and identical DMS and KMnO4 footprints. We suggest that each of these apparently unusual nucleosomes, with its corresponding adjacent non-nucleosomal region, constitutes a structural and functional unit important for the replication of the rDNA minichromosome.

We have footprinted the 5' NTS in both wild-type and maintenance mutant chromosomes. Here we report the structure of the wild-type chromosome, and focus on the unusual properties of nucleosomes 4 and 5. Results of studies with the chromosome maintenance

mutants, and an additional footprint in the non-nucleosomal region will be reported in chapter three of this thesis..

Results

Palen and Cech were the first to describe positioned nucleosomes within the 5' NTS of *Tetrahymena thermophila* rDNA (Palen and Cech, 1984). Our high resolution analysis of the C3 strain has confirmed their work, done at lower resolution in the B strain, that the majority of the 5' NTS is covered by highly positioned nucleosomes. We have used DNase I, DMS, and KMnO4 as probes of chromatin structure and protein-DNA contacts throughout the 5' NTS. These are revealed by identifying differences between the reagent induced cleavage pattern of naked DNA with that of DNA in chromatin. By using primer extension to determine sites of cleavage we have been able to map the chromatin structure and sites of potential protein-DNA interactions at single nucleotide resolution. The overall chromatin structure of this 1.9 kb region, and the approximate positions of the nucleosomes as determined from our DNase I data (see figure legend), is depicted in Figure 1b (also Figures 2a, 4a and b, 5a and b, and data not shown). In the region between nucleosomes 4 and 5, and 5 and 6, naked DNA and DNA in chromatin have the same pattern and intensity of DNase I cleavages (Figure 4a and b and 5 a and b); therefore, these sites are not covered by a positioned nucleosome. The hatched oval represents a pattern of DNase I protection and enhancements in the chromatin at the promoter region that we have described in previous work (Pan et al., 1995). Although not a component of the minimal functional

origin, genetic evidence and results of transformation with rDNA vectors indicate that a wild-type promoter is important for chromosome maintenance (this thesis; Jacek Gaertig, personal communication). Thus, as the promoter is likely to be an important component of the fully functional origin, its footprint is included in the chromatin structure of the origin region.

Nucleosome 2 is a Canonical Nucleosome - DNase I Footprinting:

Indirect end-labeling of micrococcal nuclease treated DNA provided the initial demonstration that the regions indicated in Figure 1b are bound by positioned nucleosomes (Palen and Cech, 1984). Other results with limit micrococcal nuclease treated DNA confirmed this (Cohen and Blackburn, manuscript in preparation; Gallagher and Blackburn, unpublished results). DNase I footprinting further confirms this, and demonstrates that the detectable nucleosomes are rotationally positioned to the nucleotide. This is illustrated for nucleosome 2 (Figure 2a). Nuclei from C3 strain log phase cells (+) or naked DNA (-) were incubated with the indicated concentration of DNase I for one minute at 25°C. Cleaved DNA was primer extended in the region of nucleosome 2. The positions of protein induced cleavages in chromatin are determined by identifying cleavages in the + protein lane which are absent from the - lane, or are much more intense in the + lane. Such cleavages are indicated in figure 2a for the non-coding strand of nucleosome 2 by arrows. The bracket indicates the probable extent of the nucleosome, and the asterisk indicates the probable position of the

Figure 2a: DNase I Footprinting of Nucleosome 2 on the Non-Coding Strand

0 indicates that uncleaved DNA was extended

- indicates naked DNA

+ indicates chromatin

.1 and 10 are the units of DNase I/ml in the reaction (total volume 200 microliters).

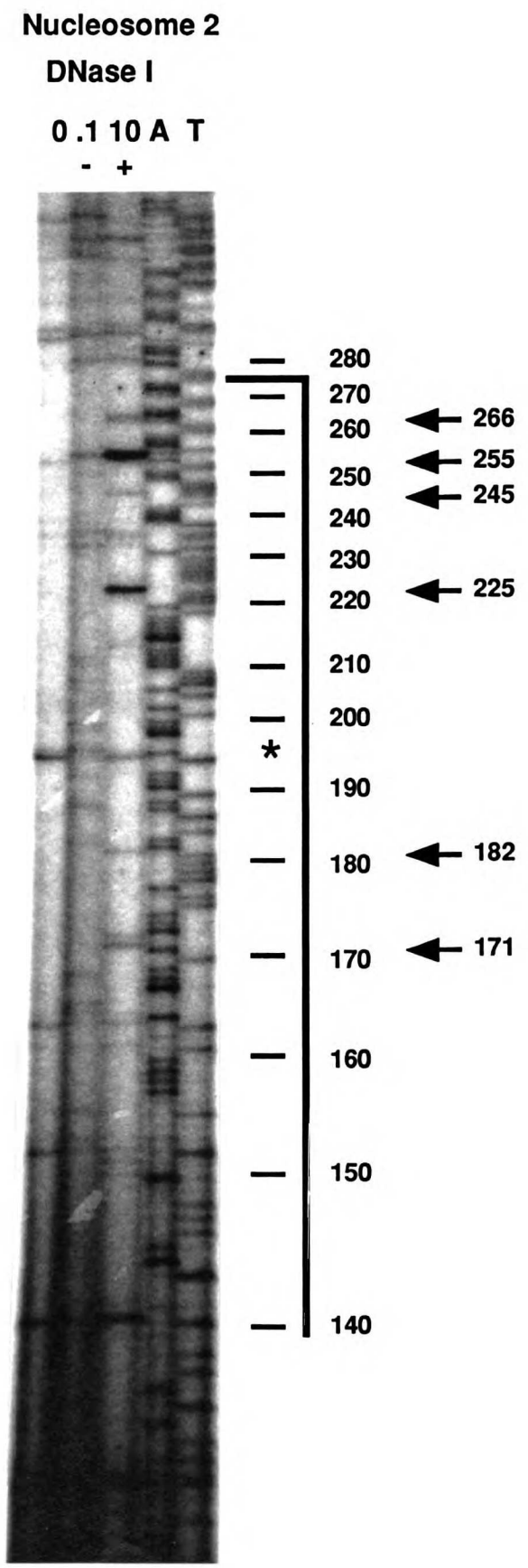
A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

The open rectangle designates the approximate extent of the nucleosome.

Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence.

Arrows point to the regions cleaved in chromatin by DNase I.

The asterisk indicates the approximate position of the nucleosomal dyad, inferred from the positions of cleavages and the presence of linker DNA.


Nucleosome 2
DNase I
0 .1 10 A T
- +
280
270
260
250
240
230
220
210
200
*
190
180
170
160
150
140
266
255
245
225
182
171

nucleosomal dyad, both have been derived from this data. The characteristic ten base pair ladder produced by DNase I on nucleosomal DNA has been well documented (Lutter, 1978; Simpson and Stafford, 1983; Simpson and Whitlock, 1976). DNase I prefers to cleave DNA at a narrow minor groove, which occurs on the surface of a nucleosome every 10 base pairs. The intrinsic asymmetries of the nucleosome result in an uneven distribution of DNase I cleavages and cleavage intensities. This was first documented by Simpson and Whitlock using isolated HeLa cell nucleosomes. They describe strong sites of cleavage at 20, 40, 50, 100, 120 and 130 nucleotides from the 5' end of the DNA. Sites at 30, 80 and 100 bases from the 5' end were rarely cut (Simpson and Whitlock, 1976). Nucleosomes within the 5' NTS of the rDNA show this same pattern of cleavages, on both strands. This is illustrated for the non-coding strand of nucleosome 2. Arrows point to strong sites of DNase I cleavage at nucleotides 255 and 225. These are likely to correspond to the susceptible positions at 20 nucleotides and 50 nucleotides from the 5' end of the nucleosome described by Simpson and Whitlock. We have assumed this in assigning the location of the dyad and the translational position of the nucleosome. Note the absence of cleavages at 265 and 195, which should correspond to positions 30 and 80 nucleotides from the 5' end, respectively. This is further evidence that these sites correspond to those positions of the nucleosome on the DNA, and that the bulk of the nucleosomes are precisely translationally positioned. The determination of the translational position is supported by the observation that DNA between nucleotides 283 and 300 is not footprinted by DNase I. This is the unprotected linker DNA

between nucleosomes 2 and 3. Note that it is possible to see the pattern of DNase I cleavage sites shift by two to three base pairs across the dyad axis. This shift is predicted from the structure of the nucleosome. There is a change in helical pitch from 10.0 base pairs per turn to 10.7 base pairs per turn for the three helical turns at the dyad, resulting in an overall shift of DNase I cleavage sites of two base pairs across the dyad (Hayes et al., 1991).

Nucleosomes 3, 4 and 5 showed strikingly similar patterns of DNase I cleavage sites (Figures 4a and a and 5a and b and data not shown). The DNA covered by nucleosomes 6 and 7 consists of 14 tandem copies of a 20 to 22 base pair repeat, and the underlying sequence regularity of the DNA dominates the pattern of the DNase I cleavages (data not shown). DNase I data was not obtained for nucleosome 1, precluding the determination of the border of this nucleosome which spans the middle of the palindrome (Figure 1b). The DNase I footprints of nucleosomes 2, 3, 4, and 5 confirm the earlier identification of these as positioned nucleosomes (Palen and Cech, 1984). Our results extend that analysis by demonstrating clearly that these nucleosomes are rotationally positioned to the nucleotide. The pattern of cleavage sites and cleavage intensities indicates that these nucleosomes are also translationally positioned to the nucleotide, because they correspond to the pattern identified by Simpson and Whitlock. These footprints are highly reproducible. The same pattern of DNase I cleavages has been identified in B strain macronuclear histone H1 knockout cells (nucleosomes 3 and 4), starved cells (nucleosome 2), and cells transformed with an rDNA

vector that with multiple tandem copies of the 5' NTS (nucleosome 5).

Nucleosome 2 is a Canonical Nucleosome: - DMS Footprinting:

In order to identify potential sites of DNA unwinding in the origin region, and to determine whether additional factors might be bound to the regions occupied by positioned nucleosomes, we chose to footprint nuclei with DMS. It has been demonstrated experimentally that DMS produces no significant footprint on nucleosomal DNA (Cartwright and Kelly, 1991; McGhee and Felsenfeld, 1979). In double stranded DNA, DMS methylates the N-7 position of G residues, which falls in the major groove, and, to a lesser extent, the N-3 position of A residues, located in the minor groove (Maxam and Gilbert, 1977). In addition, DMS methylates single stranded A residues at the N-1 position and single-stranded C residues at the N-3 position (these are base-paired in double stranded DNA). DNA remains double stranded as it wraps around the nucleosome, so no single stranded As or Cs are detected, and association with histones does not significantly enhance or reduces the extent of DNA methylation at G residues. Thus, DMS footprinting of nucleosomal DNA produces a G ladder which is of equal intensity in naked DNA and chromatin, there is no significant footprint. This characteristic absence of a DMS footprint on nucleosomal DNA is shown in Figure 2b for the coding strand of nucleosome 2. Nuclei (+) or naked DNA (-) were incubated for the indicated times with 10 mM DMS at 25°C. DNA was cleaved at sites of methylation with pyrollidine and the positions of

Figure 2b: DMS Footprinting of Nucleosome 2 on the Coding Strand

- indicates naked DNA

+ indicates chromatin

0- indicates samples quenched at 0 time and not treated with DMS

0+ indicates samples quenched at 0 time prior to DMS addition

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

The open rectangle designates the approximate extent of the nucleosome, the upper border is dotted because it has not been determined.

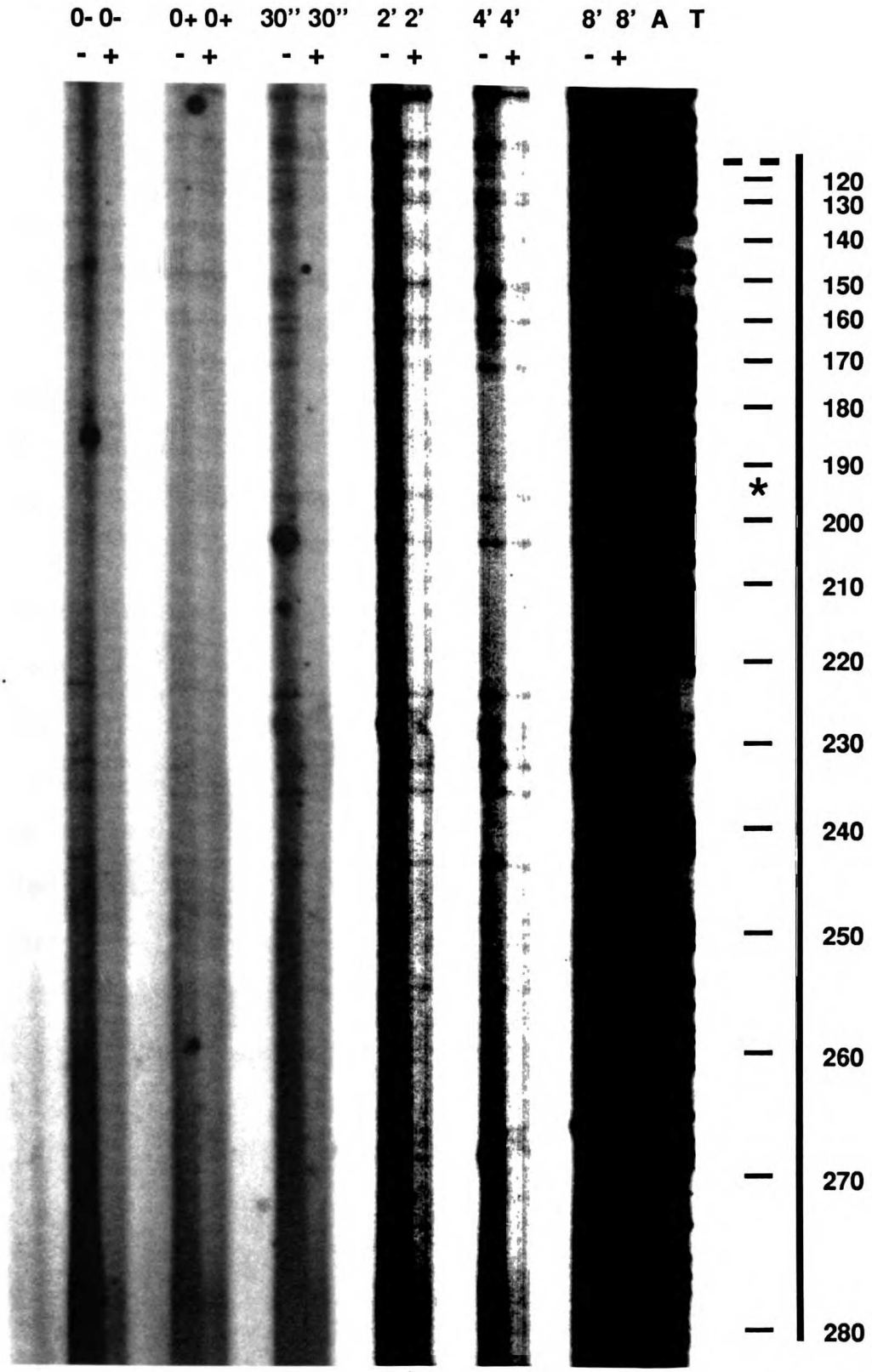
Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence.

There are no sites in chromatin that have been methylated to a greater extent than in naked DNA.

The asterisk indicates the approximate position of the nucleosomal dyad (center of the DNA wrapped around it), inferred from the positions of DNase I cleavages and the presence of linker DNA.

Nucleosome 2

DMS



cleavages determined by primer extension. The approximate extent of the nucleosome, determined from the DNase I data of Figure 2a, is indicated by the bracket. The dotted line indicates that the upstream border of the nucleosome has not been determined. Numbers along the side are the nucleotide numbers of the rDNA. The absence of arrows indicates that there are no sites which are enhanced or protected in chromatin relative to naked DNA. This is consistent with the identification of this region as nucleosomal.

Nucleosome 2 is a Canonical Nucleosome - KMnO4 Footprinting:

We chose to use KMnO4 to footprint the origin region because it detects unstacked or single stranded pyrimidines (T>>C) (Hayatsu and Ukita, 1967). Stacking refers to the hydrophobic interactions that occur between adjacent aromatic aromatic bases when the rings are flat with respect to one another. If DNA is distorted stacking will be lost, and KMnO4 is used to detect this. KMnO4 has been used to examine open complex formation at promoters (Sasse-Dwight and Gralla, 1990; Zhang and Gralla, 1989) and unwound or distorted DNA at a bacterial plasmid and a viral replication origin (Frappier and O'Donnell, 1992; Mukhopadhyay et al., 1993). The footprint of KMnO4 on nucleosomal DNA has not been described previously. Results with other reagents predict that KMnO4 might oxidize pyrimidines at 15 nucleotides and 35 nucleotides off the dyad axis, where DNA bases are unstacked (Hayes, 1995; Hogan et al., 1987; Pruss et al., 1994). However, we have found that KMnO4 produces a regular ladder of cleavages of fairly even intensity on nucleosomal DNA. This is shown for nucleosome 2 on the coding

strand in Figure 2c. Nuclei (+) or naked DNA (-) were incubated for the indicated times with 10 mM KMnO4 at 25°C. DNA was cleaved at sites of oxidation with sodium hydroxide (Gilson et al., 1993) and the positions of cleavages determined by primer extension. Arrows indicate the cleavages that are seen in chromatin and not in naked DNA, or that are more intense in chromatin than in naked DNA. These are roughly ten base pairs apart and are offset from the sites of DNase I cleavage by five base pairs (compare figures 2a and 2c). This suggests that bases are unstacked at regular intervals along the nucleosome, at sites offset by five base pairs from the sites at which the minor groove is exposed. We have also found this for nucleosome 1, coding strand; nucleosome 3 coding strand; and nucleosome 4, non-coding strand (figure 5c and data not shown). Note the absence of strong cleavages on either side of the dyad. As indicated by the sequencing ladder on the side, *Tetrahymena* DNA is very AT rich, and there are T residues at 3.5 turns off the dyad upstream and 1.5 turns off the dyad downstream that could be oxidized by KMnO4. Bases are unstacked at TA dinucleotides and this contributes to the regular pattern seen here.

Origin Proximal Nucleosomes 4 and 5 Have Typical DNase I Footprints but Unusual DMS and $KMnO_4$ Footprints:

Figures 3 a and b illustrate that the DNA in chromatin within origin proximal nucleosomes 4 and 5 is cleaved by DNase I at 10 base pair intervals, supporting the evidence from micrococcal nuclease that these regions are nucleosomal (Palen and Cech, 1984)

Figure 2c: KMnO4 Footprinting of Nucleosome 2 on the Coding Strand

- indicates naked DNA

+ indicates chromatin

0- indicates samples quenched at 0 time and not treated with KMnO4

0+ indicates samples quenched at 0 time prior to KMnO4 addition

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

The open rectangle designates the approximate extent of the nucleosome, the upper border is dotted because it has not been determined.

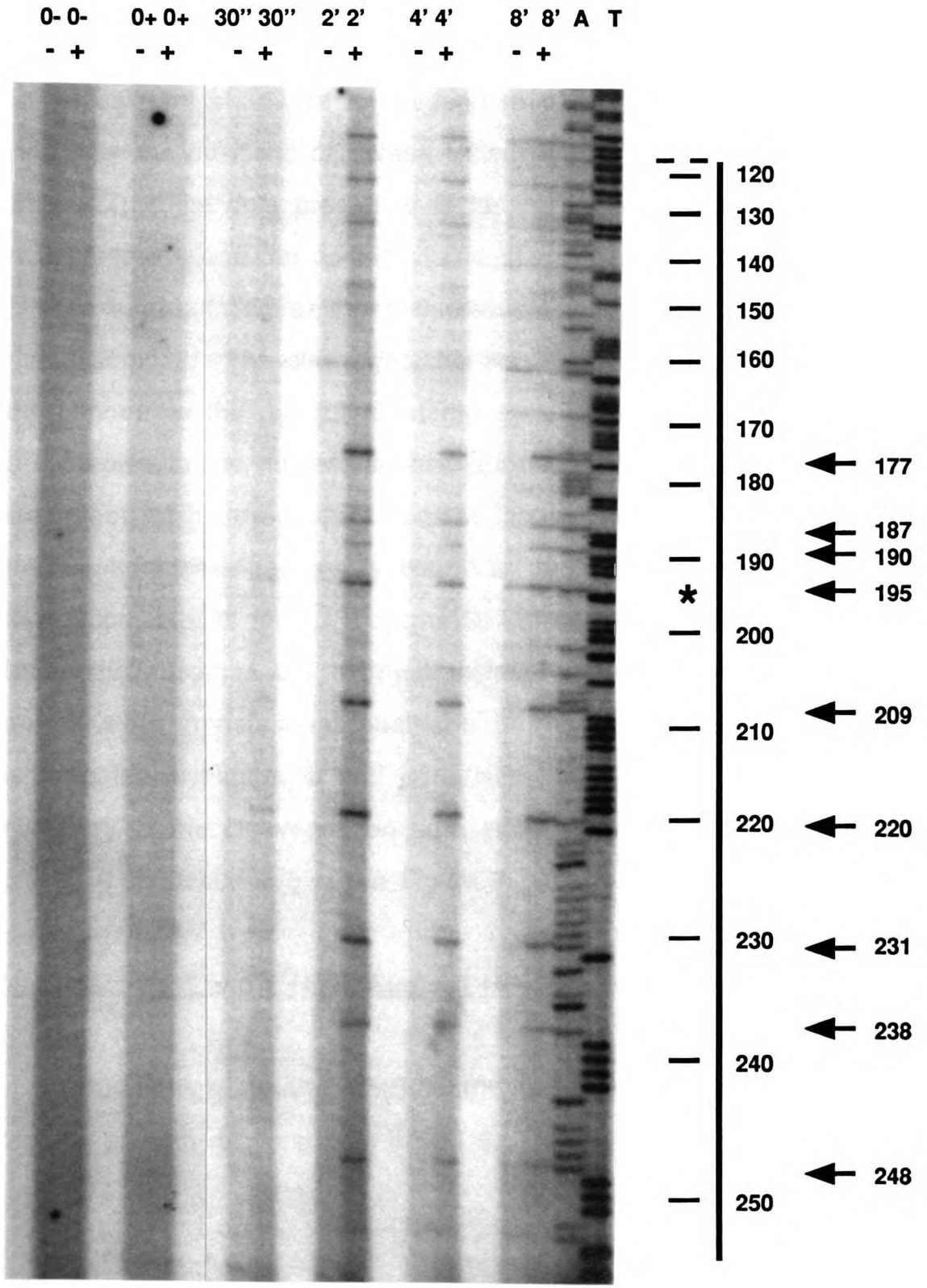
Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence.

Arrows point to the regions oxidized in chromatin by KMnO4.

The asterisk indicates the approximate position of the nucleosomal dyad (center of the DNA wrapped around it), inferred from the positions of DNase I cleavages and the presence of linker DNA.

Nucleosome 2

KMnO4



(Gallagher and Blackburn, unpublished results). The two regions are greater than 90% conserved in sequence and have virtually identical DNase I footprints. This is shown at low resolution for the coding strand of both nucleosomes in Figure 3a and b. Compare the + lanes at 10 units/ml of DNase I. At high resolution we are able to identify precisely the positions of DNase I cleavages (Figure 4a and b and Figure 7). The data presented is for the non-coding strand of nucleosome 4 and the coding strand of nucleosome 5. The results of DNase I footprinting of these nucleosomes are similar to those of nucleosome 2. The cleavages are not of uniform intensity and correspond to the pattern identified by Simpson and Whitlock, they are precise to the nucleotide, and across the dyad axis the pattern of cleavages shifts by several nucleotides . These results indicate that the DNA in these regions is bound by highly rotationally positioned nucleosomes. As for nucleosome 2, the data also suggests that these nucleosomes are highly translationally positioned. The approximate translational positions of the footprinted nucleosomes is indicated in Figure 1b and is derived from this and other data. A summary of the cleavages on both strands and their superposition on the DNA sequence is given in Figure 7.

Nucleosomes 4 and 5 Have Strong DMS Reactive Sites on the Coding Strand:

Surprisingly, strong DMS footprints were identified within the regions of both nucleosomes 4 and 5. The corresponding nucleotides

Figures 3a and b: Low Resolution View of Nucleosomes 4 and 5 Footprinted with DNase I on the Coding Strand

0 indicates that uncleaved DNA was extended

- indicates naked DNA

+ indicates chromatin

.1 and 10 are the units of DNase I/ml in the reaction (total volume 200 microliters).

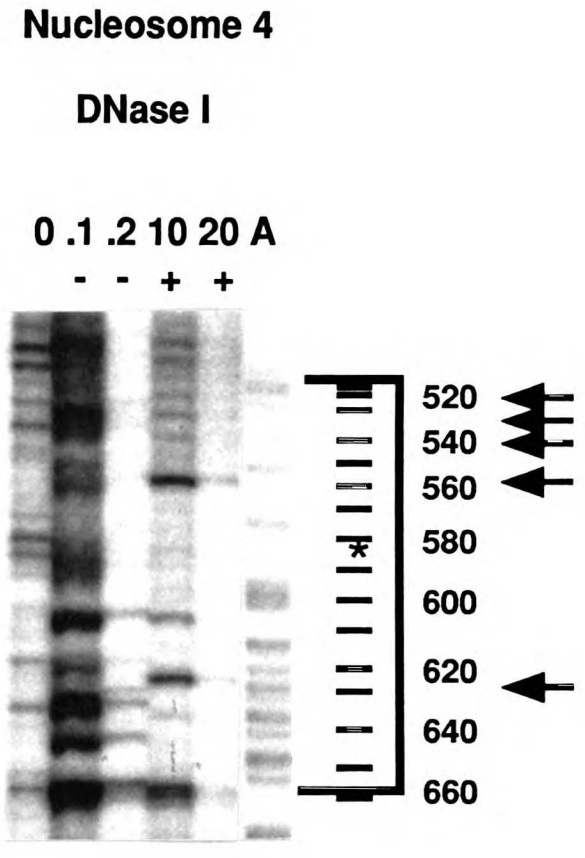
A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

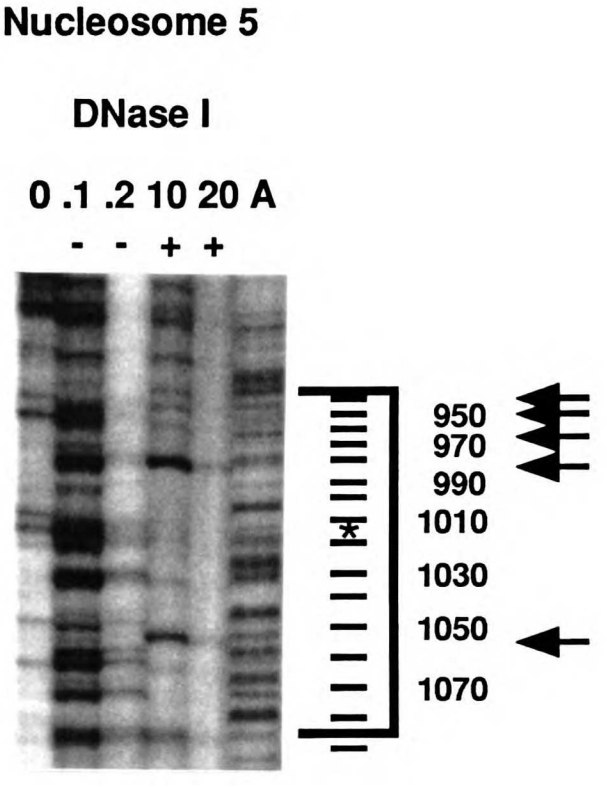
The open rectangle designates the approximate extent of the nucleosome.

Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence.

Arrows point to the regions cleaved in chromatin by DNase I.

The asterisk indicates the approximate position of the nucleosomal dyad, inferred from the positions of cleavages and the presence of linker DNA.





Figures 4a and b: High Resolution View of Nucleosome 4 on the Non-Coding Strand and of Nucleosome 5 on the Coding Strand Footprinted with DNase I

0 indicates that uncleaved DNA was extended

- indicates naked DNA

+ indicates chromatin

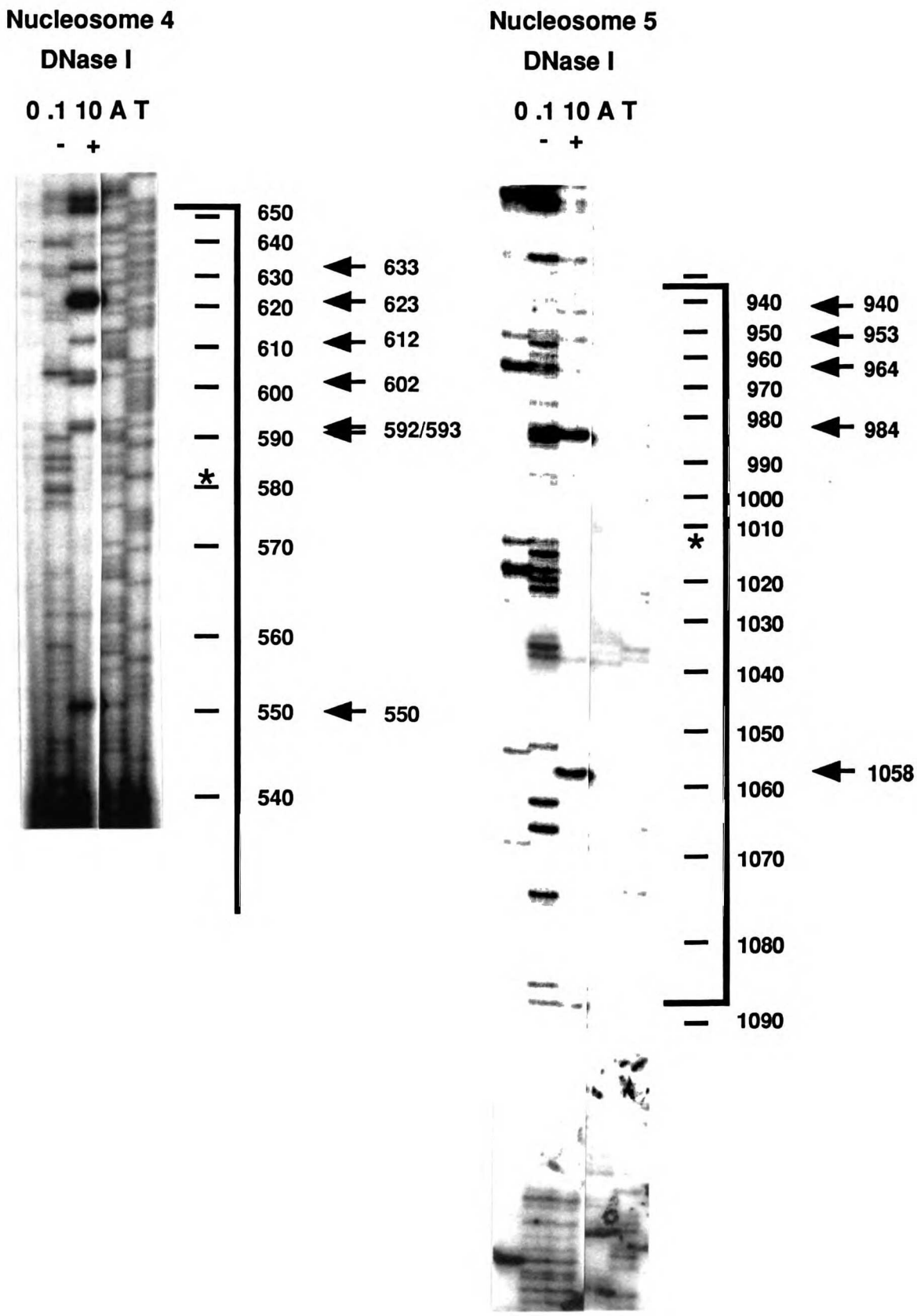
.1 and 10 are the units of DNase I/ml in the reaction (total volume 200 microliters).

A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

The rectangles designate the approximate extent of the nucleosome. Numbers beside these give the nucleotide sequence number according to the Engberg rDNA sequence.

Arrows point to the regions cleaved in chromatin by DNase I.

The asterisk indicates the approximate position of the nucleosomal dyad, inferred from the positions of cleavages and the presence of linker DNA.


Nucleosome 4
DNase I
0 .1 10 A T
- +
650
640
630
620
610
600
590
580
570
560
550
540
633
623
612
602
592/593
550
Nucleosome 5
DNase I
0 .1 10 A T
- +
940
950
960
970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
940
953
964
984
1058

are footprinted in these two homologous domains (figure 5a and b and figure 7). On the non-coding strand in each region there are three A residues which are reactive to DMS. This contrasts both with the results for nucleosome 2, and with all previous DMS footprints of nucleosomal DNA. The reactive As are identified here most clearly for the nucleosome 5 region, and are indicated by the arrows in figure 5. Compare the - and + lanes throughout the time course. Seen clearly and more intensely with increasing time of incubation with DMS are several reactive A residues upstream of the dyad (three are seen at high resolution). The exact nucleotide position is given by the sequencing ladder. Note that a sequencing reaction which terminates at dideoxy T indicates that there was an A on the opposite strand. Primer extensions of cleaved DNA terminate when the template strand ends. Thus, the run of T's of the sequencing ladder opposite the DMS reactive sites indicate that the primer extension of cleaved DNA terminated at A residues. The additional footprinted sites in Figure 5 are DMS enhanced Gs located within the three Type III repeats of Repeat 1. Type III repeats are the sites of action of Topo I on the DNA, the sites are seen in the + lanes near the top of the lanes (Bonven et al., 1985; Challoner et al., 1985). The footprinted Gs are two nucleotides downstream from the Topo I cleavage site, on the same strand. The DMS reactive As in the nucleosome 4 region (data not shown) correspond exactly to the

Figure 5: DMS Footprinting of Nucleosome 5 on the Coding Strand

- indicates naked DNA

+ indicates chromatin

0- indicates samples quenched at 0 time and not treated with DMS

0+ indicates samples quenched at 0 time prior to DMS addition

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

The open rectangle designates the approximate extent of the nucleosome, the upper border is dotted because it has not been determined.

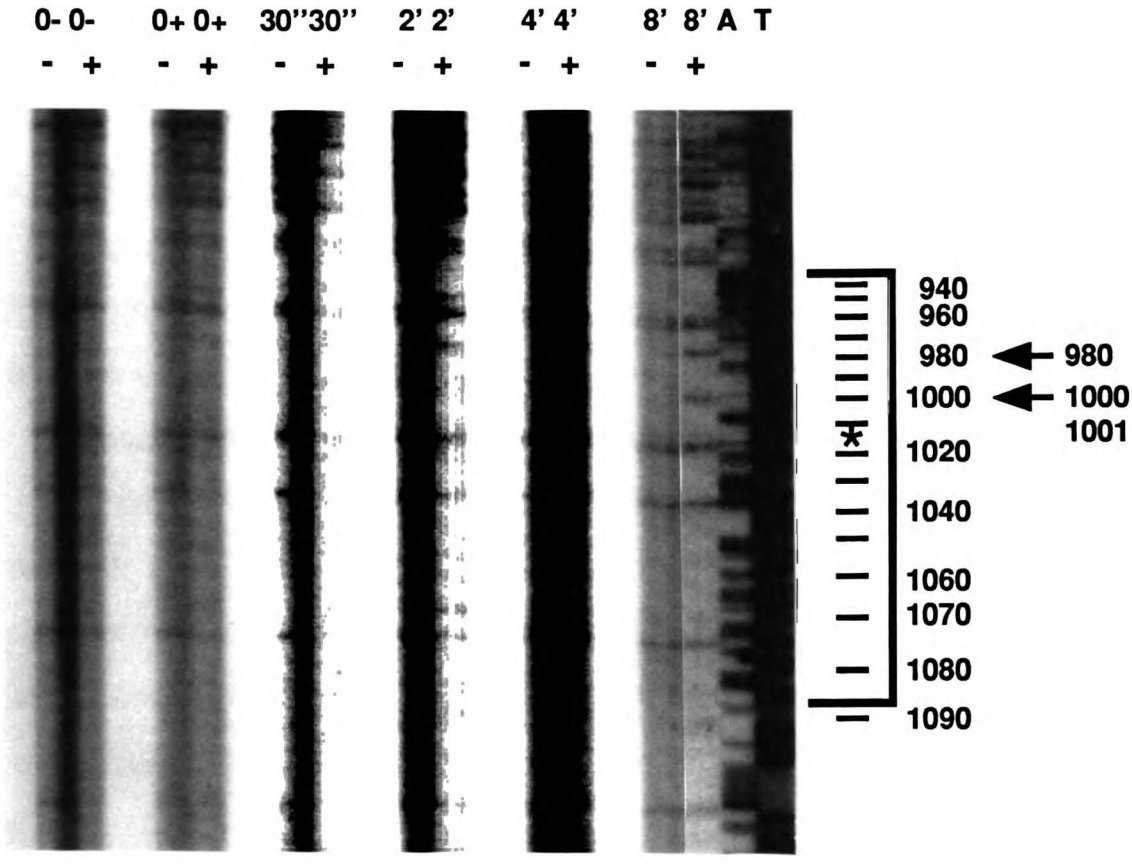
Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence.

There are no sites in chromatin that have been methylated to a greater extent than in naked DNA.

The asterisk indicates the approximate position of the nucleosomal dyad (center of the DNA wrapped around it), inferred from the positions of DNase I cleavages and the presence of linker DNA.

Nucleosome 5

DMS



positions of the footprinted As in nucleosome 5 (Figure 7). We were not able to determine whether the A residues were footprinted at the N-1 (single-stranded DNA) or N-3 (double-stranded DNA) position, since heating methylated DNA prior to primer extension resulted in a G cleavage ladder in the absence of pyrrolidine treatment. The corresponding T residues were not KMnO4 footprinted (Figure 7 and below). These sites are also present in all cell types examined including C3-*rmm8* cells and B strain H1 knockout cells (data not shown).

Nucleosomes 4 and 5 have Strong KMnO4 Sites on the Coding Strand:

Figure 6a shows the pattern of KMnO4 reactivities of the nucleosome 4 region on the non-coding strand. These are similar to the sites seen in figure 2c for the coding strand of nucleosome 2 and for nucleosomes 3, and 1 (data not shown). The cleavages are spaced approximately 10 base pairs apart, are of equivalent intensity, and are offset from the sites of DNase I cleavage by about 5 base pairs (see figure 4a and figure 7). Note that on this strand there are no other KMnO4 reactivities downstream of nucleosome 4, until the homologous KMnO4 cleavage sites of nucleosome 5 are seen high in the gel.

In contrast, on the coding strand of both nucleosomes 4 and 5, a different pattern of KMnO4 cleavages is seen. In figure 6b much of the Repeat 1 region is shown. The location of nucleosome 4 is indicated by the bracket near the top of the figure. In each of the +

Figure 6a: KMnO4 Footprinting of Nucleosome 4 on the Non-Coding Strand

- indicates naked DNA

+ indicates chromatin

0- indicates samples quenched at 0 time and not treated with KMnO4

0+ indicates samples quenched at 0 time prior to KMnO4 addition

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

The open rectangle designates the approximate extent of the nucleosome, the upper border is dotted because is has not been determined.

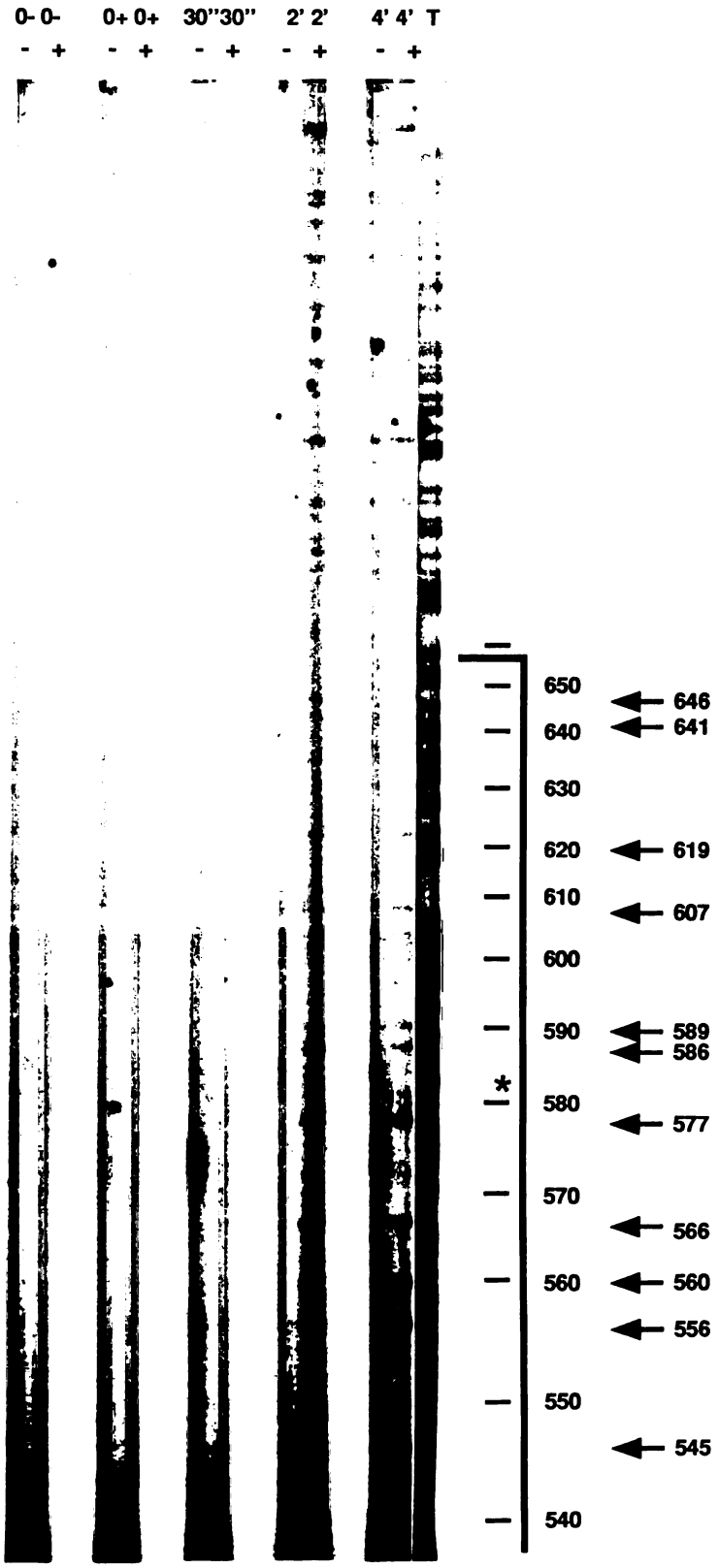
Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence (Engberg and Nielsen, 1990).

Arrows point to the regions oxidized in chromatin by KMnO4.

The asterisk indicates the approximate position of the nucleosomal dyad (center of the DNA wrapped around it), inferred from the positions of DNase I cleavages and the presence of linker DNA.

Nucleosome 4

$KMnO_4$



lanes are three strong cleavage sites not found in naked DNA. These cleavages are not at the predicted sites of unstacked bases at 1.5 and 3.5 turns off the center off the center of the nucleosome. There is no regular ladder of cleavages, as seen on the opposite strand, superimposed on this pattern.

On the coding strand there are several KMnO4 reactive sites between nucleosomes 4 and 5 (proceeding down toward the bottom of the gel, Figure 6b); however, the relative intensities of cleavages at sites not arrowed has varied between different strains and preparations and we cannot be sure that there is a difference between chromatin and naked DNA. There is a KMnO4 footprint at the Type IIIa repeat which was also footprinted on this strand with DMS (three strong sites in + lanes above nucleosome 5, Figure 5a). The footprinted residue on this strand is that at which Topo I cleaves the DNA, so that a cleavage is seen in the 0 timepoints in naked DNA and chromatin. However, in the presence of KMnO4 there is a greatly enhanced cleavage of this residue in chromatin.

Figure 6c illustrates that chromatin of the nucleosome 5 region, bracketed from nucleotide 935 to 1087, is cleaved the at the residues which correspond to those cleaved by KMnO4 in nucleosome 4 (figure 5b and figure 7). The footprint of nucleosome 4 on this strand is seen high in the figure. The footprints within the Type III repeats of Repeat 1 are seen below this (compare with figure 5b). These nucleosomal KMnO4 reactive sites are found on the same

Figures 6b and c: KMnO4 Footprinting of Nuclesomes 4 and 5 on the Coding Strand

- indicates naked DNA

+ indicates chromatin

0- indicates samples quenched at 0 time and not treated with KMnO4

0+ indicates samples quenched at 0 time prior to KMnO4 addition

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

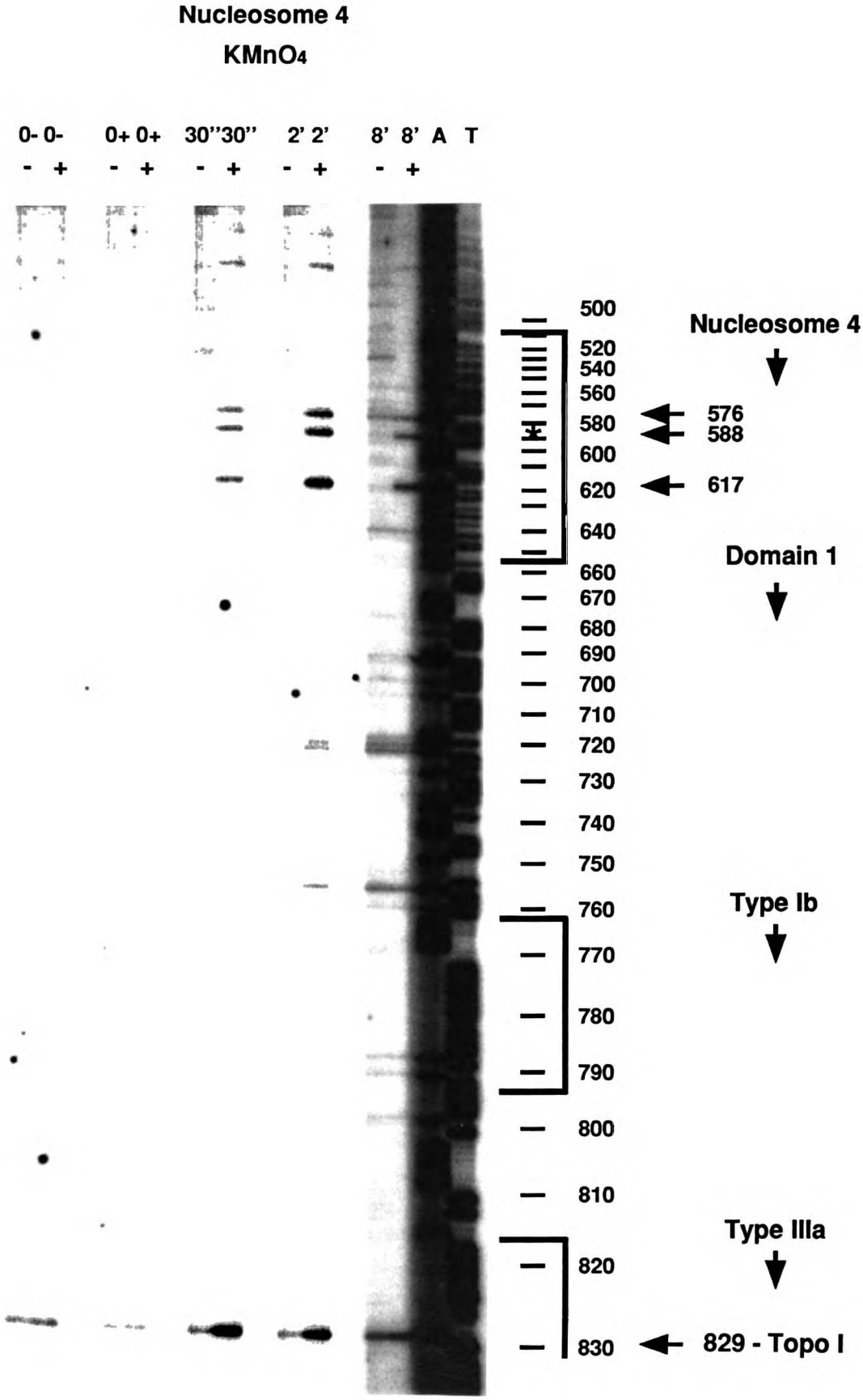
A and T are sequencing lanes. Plasmid DNA was primer extended in the presence of ddA or ddT.

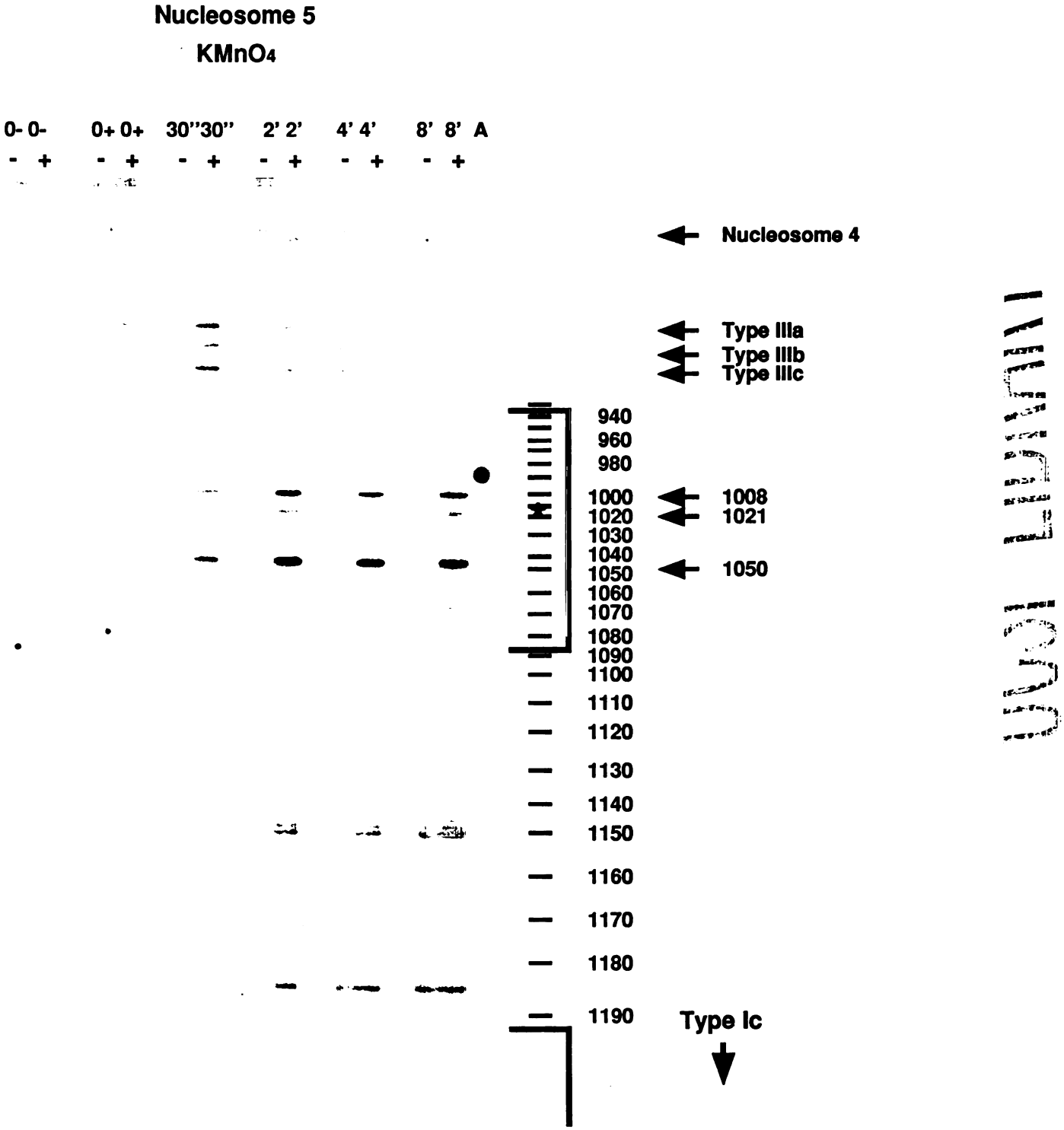
The open rectangle designates the approximate extent of the nucleosome, the upper border is dotted because is has not been determined.

Numbers beside this give the nucleotide sequence number according to the Engberg rDNA sequence (Engberg and Nielsen, 1990).

Arrows point to the regions oxidized in chromatin by KMnO4.

The asterisk indicates the approximate position of the nucleosomal dyad (center of the DNA wrapped around it), inferred from the positions of DNase I cleavages and the presence of linker DNA.


Nucleosome 4
KMnO4
0- 0- 0+ 0+ 30"30" 2' 2' 8' 8' A T
- + - + - + - + - +
500
520
540
560
576
580
588
600
620
617
640
660
670
680
690
700
710
720
730
740
750
760
770
780
790
800
810
820
830
829 - Topo I
Nucleosome 4
Domain 1
Type Ib
Type IIIa


Nucleosome 5
KMnO4
0- 0- 0+ 0+ 30"30" 2' 2' 4' 4' 8' 8' A
- + - + - + - + - + - +
Nucleosome 4
Type IIIa
Type IIIb
Type IIIc
940
960
980
1000
1020
1030
1040
1050
1060
1070
1080
1090
1100
1110
1120
1130
1140
1150
1160
1170
1180
1190
1008
1021
1050
Type Ic

strand as that on which the DMS reactive A residues are found. The data for the three reagents on the two nucleosomes is summarized in Figure 7 and is discussed below. Figure 8 illustrates where on the surface of the nucleosome these residues are likely to be located.

Discussion

The importance of the role of chromatin structure in the regulation and function of numerous biological processes has just begun to be illuminated and is likely to be a rich field of scientific exploration for many years (Felsenfeld, 1992; Wolffe, 1991). No longer are nucleosomes viewed as mere packaging agents which bind DNA non-specifically and have a purely structural role in the cell. It is now clear that the interaction of nucleosomes with DNA can be sequence specific, and that the placement of nucleosomes on specific sequences can have important consequences, for example, the activation or repression of transcription (Wolffe and Dimitrov, 1993).

To date the role of chromatin structure has been studied most fully in the regulation of transcription. The observations that have been made include the specific positioning of nucleosomes on, or off, regulatory elements, chromatin remodeling during transcriptional activation, the differential expression during development of variant linker histones and the correlation of this with a change in gene expression of somatic and maternal 5S genes, and the identification of core histone variants that may be specifically associated with transcribed DNA (Allis et al., 1980; Prioleau et al., 1994; Schild et al.,

1993; Wolffe and Dimitrov, 1993). Other important results are the identification of histone modifications and their role in transcriptional activation and repression, and the identification of complexes which remodel chromatin (Orlando and Paro, 1995; Wolffe, 1994).

It is likely that chromatin structure will play an important role in any process involving DNA within the eukaryotic cell, transcription, recombination, repair, segregation, replication. Except for the identification of highly positioned nucleosomes in the vicinity of some well studied origins, ARS1 and the *Tetrahymena* rDNA, nucleosomes have not been shown to have a distinctive structure or role in DNA replication. Here we have described in detail the chromatin structure of an origin region. We show that the nucleosomes in the vicinity of the origin are highly rotationally positioned and are likely to be highly translationally positioned, and we show that origin proximal nucleosomes in the *Tetrahymena* rDNA produce a distinctive footprint. In chapter three we suggest that one role for these positioned nucleosomes may be to facilitate the juxtaposition of complexes at the promoter region and origin in order to enhance replication.

There are several possible explanations for the unusual footprints identified in the nucleosome 4 and 5 regions. One is that a sub-population of rDNA minichromosomes is not nucleosomal in these regions, but is bound by another factor(s) that induces the footprint. It is possible that another factor has bound nucleosomal

DNA directly. A third explanation is that the structure of the nucleosome itself has been altered, through the incorporation of a variant histone or by the modification of histones within the nucleosome. Based on our footprints of other 5' NTS nucleosomes, if there is a population of molecules which is apparently nucleosomal on one strand we should see the regular nucleosomal pattern on the opposite stand. Because this is not seen, we suggest that the DMS and the KMnO4 footprints that we observe in the nucleosome 4 and 5 regions are due to a modification of the nucleosome itself, either intrinsically or extrinsically. This could happen through a modification of the histones, through the substitution of a histone protein with a histone variant, or through the binding of another protein(s) to the nucleosome. It will be of interest to determine what has caused this unusual footprint, what function it serves, and whether similar modifications are found in other origin proximal nucleosomes.

Figure 7 summarizes the data presented in the paper concerning the cleavages with the three reagents on nucleosomes 4 and 5. The sequence differences between the nucleosome 4 and 5 regions are bolded above the main sequence. In this summary it can be seen that the two nucleosomes have very similar footprints. Figure 8 is a representation of a nucleosome and includes a suggestion for the placement of the observed DMS and KMnO4 sites on the surface of the nucleosome based on the rotational information provided by the DNase I data. It is important to note that the KMnO4 sites are not integral multiples of 10 base pairs away from each other. This

indicates that if the KMnO4 sites are due to protein binding or to an altered structure on the surface of a nucleosome, that nucleosome has only one translational position. This supports our interpretation of the DNase I data.

The chromatin structure of the *Tetrahymena* rDNA origin is very similar to that of the TRP1ARS1 plasmid of *S. cerevisiae*. On this 1.4 kb replicon there are seven highly positioned nucleosomes and two nucleosome free regions (Thoma et al., 1984). One is located at the 5' end of the TRP1 transcription unit, the other is at the 3' end and contains the A consensus ARS element which binds the origin recognition complex (ORC), and adjacent B element, also critical for origin function (Simpson, 1990). In addition, the work which defined the functional maintenance elements of the TRP1ARS1 plasmid also identified the C element. Deletion of C has no effect on origin function in the presence of an intact B element. However, the C is required for origin function in the absence of the B element (Celniker et al., 1984). The location of the C element corresponds to the region of the positioned nucleosome downstream of the A element. It is possible that the origin proximal nucleosomes 4 and 5 of the *Tetrahymena* rDNA 5' NTS are analogous to the C element of ARS1, although the C element has not been well defined. We suggest that the unusual footprinting properties of nucleosomes 4 and 5, and their precise inclusion in the duplication of Repeats 1 and 2, indicate that they are part of a structural and functional unit important for wild-type origin function. There is at least one

Figure 7: Sequence Alignment and Cleavage Sites

The data presented in Figures 4, 5, and 6 is summarized in this figure.

The sequence of the first 160 nucleotides of Repeat 1 corresponds to the position of Nucleosome 4, and is written out in plain text. The top stand is the coding strand (this refers to the rRNA genes located 1100 base pairs away).

The sequence of Repeat 2 corresponds to the position of Nucleosome 5 and is highly homologous to that of Repeat 1 (Nucleosome 4).

Where the sequence differs the nucleotide of Repeat 2 has been placed above that of Repeat 1 for the coding strand and below that of Repeat 1 for the non-coding strand in bold.

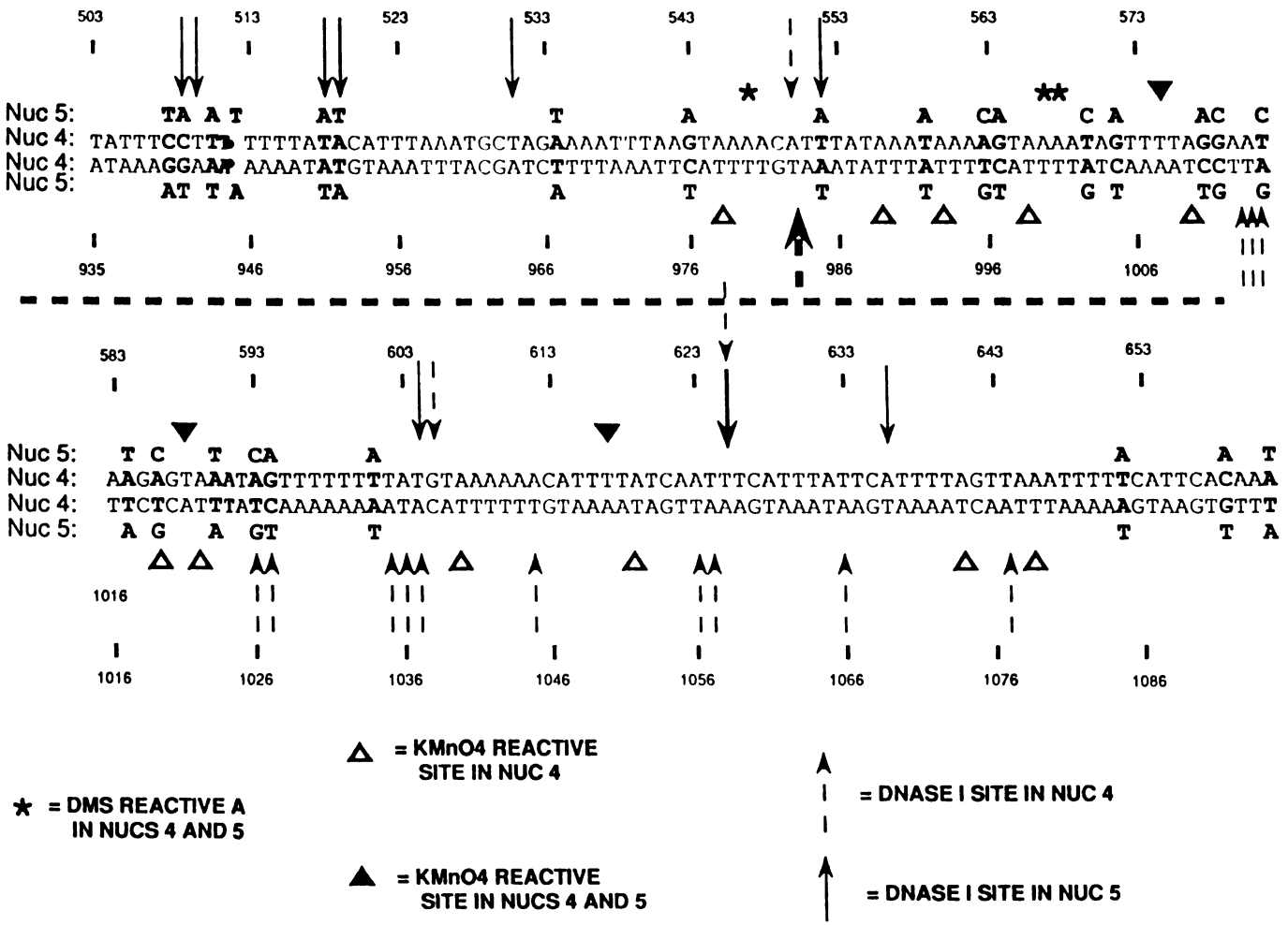
The space indicates a gap in Repeat 1.

Long arrows are DNase I cleavages, dashed are nucleosome 4 and solid are nucleosome 5.

Asterisks are DMS reactive sites present on both nucleosomes 4 and 5.

Solid triangles are $KMnO_4$ reactive sites present on both nucleosomes 4 and 5, hollow triangles are $KMnO_4$ reactive sites present on nucleosome 4.

SEQUENCE ALIGNMENT AND SUMMARY OF CLEAVAGE SITES:



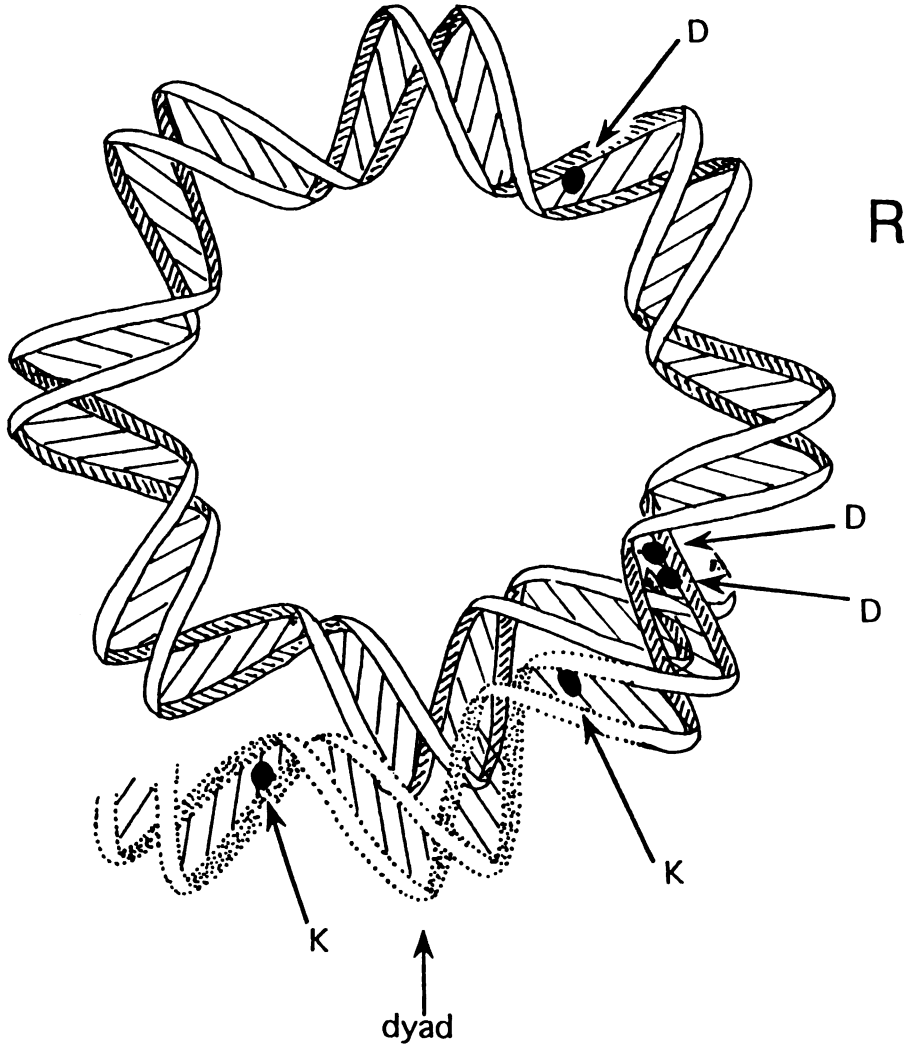
example of the incorporation of an unusual histone at a specialized region - the centromere (Wolffe, 1995). The specialized histone at the centromere is a variant histone H3. Perhaps there are specialized or modified histones at replication origins. Three histone H3 genes have been identified in *Tetrahymena*, two encode the same DNA sequence, and the third, a variant, is expressed in both growing and non-growing cells (Gorovsky, personal comunication). Perhaps there is a modification to a histone rather than a histone variant.

Alternatively, the unusual footprints of nucleosomes 4 and 5 may reflect the presence of a bound factor. Candidate proteins for this factor are the high mobility group (HMG)-like proteins found in *Tetrahymena.* HMG proteins were originally described in higher eukaryotes and are the most abundant class of non-histone chromosome associated proteins. They are identified by their physical properties and are grouped into three classes, HMG 1/2, HMG 14/17, HMG I/Y (Bustin et al., 1990). As a group they have been implicated in transcription, replication, and repair. HMG 14/17 have been shown to bind directly to nucleosomes (Alfonso et al., 1994), HMG 1/2 have been shown to bend DNA (Paull et al., 1993), and HMG I/Y are associated with H1 depleted chromatin (Zhao et al., 1993). The sequence specific HMG proteins SRY and TCF-1 bind the consensus sequence A/TA/TCAAAG and footprinting has demonstrated that methylation of any of these A residues at the N-3 position interferes with binding of these proteins. This, and other evidence, indicates that these HMG proteins bind through the minor

groove (van de Wetering and Clevers, 1992). In *Tetrahymena* four HMG-like proteins have been described and two of these have been well-studied (Wang and Allis, 1993). These are called HMG B and HMG C and they have sequence similarities to the HMG 1/2 group of higher eukaryotes and physical similarities to the HMG 14/17 group (Wang and Allis, 1993). The two sites of DMS methylation in nucleosomes 4 and 5 are within the sequence G/ATAAAAT/C, at the 4th, and 4th and 5th positions, respectively. As mentioned in the results section, it was not possible to determine definitively whether these DMS As were methylated at the N-1 position, accessible in single stranded DNA only, or the N-3 position in the minor groove of double stranded DNA. However, the absence of KMnO4 footprints of the T residues on the opposite strand suggests strongly that these As are methlyated at the N-3 position and reflect an interaction of proteins with the minor groove. DMS footprinting of HMG proteins on naked DNA or nucleosomal DNA has not been reported, so we cannot compare our results to those of others. However, the observations that HMG proteins bind nucleosomal DNA, bind in the minor groove, and may be involved in replication make the *Tetrahymena* HMG-like proteins good candidates for the factor(s) causing the DMS footprints of A residues within nucleosomes 4 and 5 of the rDNA origin region described here. The recent development of gene knockout technology in *Tetrahymena* makes it possible to test the involvement of the HMG -like proteins that have been cloned in the generation of these unusual footprints (Shen et al., 1995).

Figure 8: Sites of DMS and KMnO4 Reactivities on the Coding Strand of Nucleosomes 4 and 5

The data summarized in Figure 7 have here been used to determine where the DMS and KMnO4 sties are likely to fall on the surface of the nucleosomes of the origin region.



Sites of DMS and KMnO4 Reactivities on the Coding Strand of Nucleosomes 4 and 5

K = KMnO4

D = DMS

As mentioned in the introduction, the physical and functional origins of the *Tetrahymena* rDNA minichromsome have not been precisely mapped. Both localize to a 900 base pair region within the 1.9 kb 5' NTS in which two copies of a roughly 400 base pair sequence are found in tandem direct repeat, the functional origin has been shown to include additional sequence 3' of the repeats. Other Tetrahymenid ciliates have only one copy of the repeated sequence in their rDNA and it is not clear which of the two contains the functional rDNA origin in *Tetrahymena thermophila*. It is possible that either repeat can function as an origin. This duplication has confounded the identification of the minimal origin in *Tetrahymena*, but it has provided an internal control in our footprinting analyses. By footprinting the two regions appear identical, even at regions of sequence variability. This suggests that they are functionally similar. We suggest that the nucleosomal region within the tandem repeat is an important component of the function conferred by this repeat, DNA replication. A precedent for this is found in ARS1 of *S.cerevisiae* (Celniker et al., 1984). Possible functions of these unusual nucleosomal regions include nucleosome positioning to clear the origin, binding licensing factors that restrict origin firing to once per cell cycle, altering DNA structure to facilitate DNA unwinding, and binding proteins that target replication factors to the origin region.

If the structure of the *Tetrahymena* origin is analagous to that of ARS1 the non-nucleosomal regions between nucleosomes 4 and 5 and 5 and 6 should contain A and B element homologues. However, we do

not find the DNase I footprints identified *in vivo* in *S.cerevisiae* at these sites. We do find DMS and $KMnO_4$ footprints at the well defined topoisomerase I sites (Bonven et al., 1985; Challoner et al., 1985), and a single DMS reactive A residue sixty-one nucleotides upstream of the highly conserved Type I repeat. This DMS residue is lost in a chromosome maintenance mutant (this thesis) and is the one candidate for an origin footprint. These will be reported in more detail in conjunction with the description of the footprinting of the maintenance mutants. Here we have focused on the overall structure of the origin region, and the unusual footprints of two origin proximal nucleosomes.

Materials and Methods

Cell Strains and Culture: The following *Tetrahymena thermophila* strains were used; SF137, a C3 hemizygous strain created by Jeff Kapler; C3491-1a, a C3 wild-type strain obtained from Ed Orias, 42-b, a 6G + Ter transformant (Pan et al., 1995), SB2120 (C3-*rmm1 Pmr+*); SB1934, B strain rDNA; SF108, a C3-*rmm7* hemizygous strain created by Jeff Kapler, C3-*rmm3* Rd2#1, a C3-*rmm3* homozygous strain created by Renata Gallagher; SF112, C3-*rmm8* hemizygous strain created by Jeff Kapler; H1 knockout cells with B strain rDNA, from the Gorovsky lab, two prD4-1 transformants with differing numbers of reiterations of the central origin region. Cells were grown at 30°C in 2% PPYS (2%Proteose Peptone (Difco Laboratories, Detroit , Michigan), 2% yeast extract (Difco), 0.003% Sequestrine (CIBA-GEIGY Corporation, Summit, New Jersey)) with 1 x PSF (100 U of penicillin per ml, 250 picograms of amphotericin B

per ml, and 100 micrograms of streptomycin) (GIBCO Laboratories, Grand Island, N.Y.) per ml in flasks with aeration by gentle swirling (100 rpm) on a gyratory shaker. We starved cells by pelleting them from log phase cultures in a clinical centrifuge at 2 k rpm for two minutes in 50 ml Falcon tubes at room temperature, washing them twice in 10 mM Tris-HCl, pH 7.5 or 1 x Dryl's with Calcium, and resuspending them in 1 x Dryl's with calcium plus 1 x PSF. Cells were refed after starvation with 5% PPYS to a final concentration of 1 or 2 % PPYS.

Nuclei Prep Protocol: Nuclei were isolated as described previously (Budarf and Blackburn, 1986) except that nuclei from 10^7 cells were resuspended in 700 microliters of Buffer A (60 mM KCl, 15 mM NaCl, 0.5 mM spermidine, 0.15 mM spermine, 15 mM Tris-HCl (pH 7.4), 2 mM $CaCl_2$, 15 mM Beta-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride). Spermine and spermidine were eliminated in some KMnO4 experiments, this had no effect on the footprint. Beta-mercaptoethanol was omitted from Buffer A in DMS and potassium permanganate preps. Treatment of nuclei and isolated DNA with DNase I has been described previously (Pan et al., 1995). Treatment of nuclei and isolated DNA with DMS and $KMnO_4$ is described below. DMS: One hundred microliter aliquots of nuclei in Buffer A without BME, containing approximately 10^6 macronuclei or about 11 micrograms of macronuclear DNA, were preincubated at 25°C for 2 minutes. Ten microliters of DMS (10 M stock, Aldrich, diluted 1:100 in ddH_2O) was added to a final concentration of 10 mM. Samples were mixed for 10 seconds and incubated for 30 seconds, 2

minutes, four minutes, or eight minutes total. Two 0 timepoints were taken, one in which DMS was added after the addition of the quenching reagent (0+), and one in which no DMS was added (0-). DMS reactions were quenched by the addition of 900 microliters of buffer A with .3 M BME at 4°C and were placed on ice. Nuclei were pelleted in a microfuge at 4°C by a 6 minute spin at 6.5 k rpm. They were resuspended in 200 microliters of Proteinase K solution (20 mM Tris-HCl (pH 7.4), 5 mM EDTA, 1% sodium dodecyl sulfate (SDS), .5 mg/ml proteinase K) mixed 1:1 with Buffer A containing .3 M BME (final concentration of .15 M). Samples were incubated at 37°C overnight, extracted with phenol-chloroform and chloroform, and precipitated with 2.5 M sodium acetate and ethanol. DNA was resuspended in 20 microliters of double distilled water. Naked DNA was prepared from nuclei in the following manner; to each of six 100 microliter aliquots of nuclei, 100 microliters of Proteinase K solution (see above) was added. Samples were incubated at 37°C overnight, extracted with phenol-chloroform and chloroform, and precipitated with 2.5 M sodium acetate and ethanol. DNA was resuspended in 20 microliters of Buffer A. For DMS treatment of naked DNA 90 microliters of Buffer A was added to 10 microliter aliquots of DNA in Buffer A, and samples were treated with 10 mM DMS as above for nuclei except that reactions were quenched with the addition of 35 microliters of 3x DMS stop mix (7.5 M Ammonium acetate, 1 M BME, 250 micrograms per milliliter of tRNA) at 4°C, and precipitated directly with ethanol. Samples were resuspended in ddH_2O. Prior to primer extension most samples were chemically cleaved; 10 microliters of modified DNA was added to 90

microliters of pyrollidine in ddH_2O and incubated at 90°C for 15 minutes. Samples were precipitated with sodium acetate and ethanol and carrier tRNA, lyophillized twice from 100 microliters of ddH_2O, and resuspended in 20 microliters of TE (10 mM Tris-HCl (pH 7.4), 1 mM EDTA) plus RNAse A (10 micrograms per ml). In some cases DNA was not chemically cleaved, instead DNA was cleaved by incubation at 95°C for 5 to 10 minutes during the denaturation step of the primer extension reaction.

Potassium Permanganate: One hundred microliter aliquots of nuclei in Buffer A without BME, containing approximately 10^6 macronuclei or about 11 micrograms of macronuclear DNA, were preincubated at 25°C for 2 minutes. 2.7 microliters of .37 M $KMnO_4$ was added to a final concentration of 10 mM. Samples were mixed for 10 seconds and incubated for 30 seconds, 2 minutes, four minutes, or eight minutes total. Two 0 timepoints were taken, one in which $KMnO_4$ was added after the addition of the quenching reagent (0+), and one in which no $KMnO_4$ (0-). $KMnO_4$ reactions were quenched by the addition of 900 microliters of buffer A with 1 M BME at 4°C and were placed on ice. Nuclei were pelleted in a microfuge at 4°C by a 6 minute spin at 6.5 k rpm. They were resuspended in 200 microliters of Proteinase K solution (see above) with 1 M BME. Samples were incubated at 37°C overnight, extracted with phenol-chloroform and chloroform, and precipitated with 2.5 M sodium acetate and ethanol. DNA was resuspended in 100 microliters of double distilled water. Naked DNA was prepared from nuclei in the following manner; six 100 microliter aliquots of nuclei were

preincubated at 25°C for 2 minutes, then 900 microliters of Buffer A with 1 M BME at 4°C was added to each tube. Samples were spun at 6.5 K rpm for 6 minutes at 4°C in a microfuge and the nuclei pellet was resuspended in 200 microliters of Proteinase K solution (see above) with 1 M BME and placed at 37°C overnight. These were extracted with phenol-chloroform and chloroform, and precipitated with 2.5 M sodium acetate and ethanol. Samples were resuspended in 100 microliters of double distilled water. These were treated with potassium permanganate as for nuclei above except that reactions were stopped with the addition of 35 microliters of 3 x potassium permanganate stop mix (7.5 M ammonium acetate, 3 M BME, 250 micrograms per milliliter tRNA) at 4°C and were precipitated directly with ethanol. Samples were resuspended in 100 microliters of double distilled water. Prior to primer extension potassium permanganate samples were cleaved with base according to the protocol of Gasser (Gilson et al., 1993). 2 M sodium hydroxide was added to 200 mM and samples were placed at 70°C for 20 minutes. These were neutralized by the addition an equal volume of 200 mM HCl and one tenth volume 1.5 M Tris-HCl (pH 8.0) with 4 mM EDTA. Samples were precipitated with sodium acetate and ethanol and resuspended in double distilled water.

Primer Extension Reactions: Relative DNA amounts were determined by the use of a fluorometer. Primer extension reactions were carried out on equivalent amounts of DNA as described (Pan et al., 1995) except that in some of the primer extension reactions the phenol-chloroform step was omitted and DNA was precipitated on

dry ice for 10 minutes. In addition, as noted above, in some cases DMS treated DNA was heat cleaved by incubation at 95°C for 5 to 10 minutes during the denaturation step of the primer extension reaction. The following primers were used to detect cleavages on the transcribed strand:

	Sequence 5' to 3':	nt:
18	GAA TTT TAC CTT CGA AAA TCA C	51-72
19	TAC AAA TTT ACA AAT TTT CAA GC	264-286
20	AAA GCA TCT AAA AAT GGA CA	474-493
5	TTA GGA ATA IGA GTA AAT AG	575-594
60 r.c.	ATC ACT TTT TTT GAG AGT TG	798-817
Ib	AAC AAT TTT AAC AAC ATG CGT ATA TC	1001-1026
36 r.c.	CTT TTG CAA CTT TTG AGA CTT CG	N.A.
9	TGA TTT AGG AGA AAT TTT GAG	1321-1341
63	CTC GCT TAA TAT TCA GCG GAG	1494-1513
11	GCT CTA AAT TAA ATT AGA CTT AGT G	1665-1689

The following primers were used to detect cleavages on the non-transcribed strand:

	Sequence 5' to 3':	n.t.
4	CAT CTA ATT CTT ATC TAC TC	111-130
23	GCA TTA AAG TAG TCA AAT AGC	327-347
5 r.c.	CTA TTT ACT CAT ATT CCT AAA AC	572-594
60	TTC AAC TCT CAA AAA AAG TG	800-819
6	AAT GAT ATA CGC ATG CTG TTA	1009-1026
36	CAC GAA GTC TCA AAA GTT G	N.A.

21	CTC AAA ATT TCT CCT AAA TCA	1321-1341
63 r.c.	CTC CGC TGA ATA TTA AGC GAG	1494-1513
10	CGC TAT TTT TCA CTA AGT CTA	679-1699
12	TCT TAC TGA AGC TCA AAT CGA GCT G	1948-1924

C3 and B strain cells have two different alleles of rDNA. Unless noted the sequence is common to both. Where they differ the nucleotide has been bolded if is C3 specific and underlined if it is B specific.

CHAPTER 3:

Tetrahymena rDNA Mutants Indicate that Interactions Between a Promoter Region and an Origin Region are Important for Wild-Type Chromosome Maintenance

Abstract

We have characterized a cis-acting mutation in the rDNA minichromosome of *Tetrahymena* which causes a maintenance phenotype. rDNA maintenance mutant 3 (C3 - *rmm3*) rDNA is maintained in cells homozygous for this allele but is lost from the *Tetrahymena* macronucleus when it is in competition with the B rDNA allele. The base change responsible for the maintenance defect of *rmm3* is a deletion of an A residue in a run of 11 A's in the phylogenetically conserved promoter distal Type I repeat of the Pol I promoter region of the rRNA genes (Type Ic repeat). Previously described maintenance mutants (*rmm1* and *rmm4*) have base changes in a Type I repeat, located at -700 from the start site of rRNA transcription, in the replication origin region (Type Ib repeat). In order to investigate a link between replication and transcription we have analyzed promoter function, replication intermediates, and origin region and promoter region protein-DNA interactions in wild-type and *rmm3* homozygous cells. Nuclear run-ons revealed no defect in the initiation of transcription in *rmm3* cells. In contrast, homozygous *rmm3* cells do appear to have a replication phenotype. In log phase cells there is an accumulation of replicating intermediates in the 5' non-transcribed spacer which is not seen in wild-type cells. Finally, footprinting of *rmm3* cells reveals that they have lost the wild-type Pol I promoter footprint and lack a wild-type origin region footprint. We also report the base changes of two new maintenance mutants. *rmm7* has a base change in the same Type I repeat that is mutated in *rmm1* and *rmm4,* located in the origin region (Type Ib). *rmm8* has a base change at -19 from

the start site of transcription which is responsible for its maintenance defect. The location of these base changes in the new maintenance mutants supports our hypothesis that the promoter region and the origin region are both important for wild-type maintenance of the chromosome (Larson et al., 1986; Yaeger et al., 1989). The footprinting results of the *rmm3* strain suggest that there is a physical interaction between the promoter region and the origin region that is important for wild-type maintenance of the chromosome.

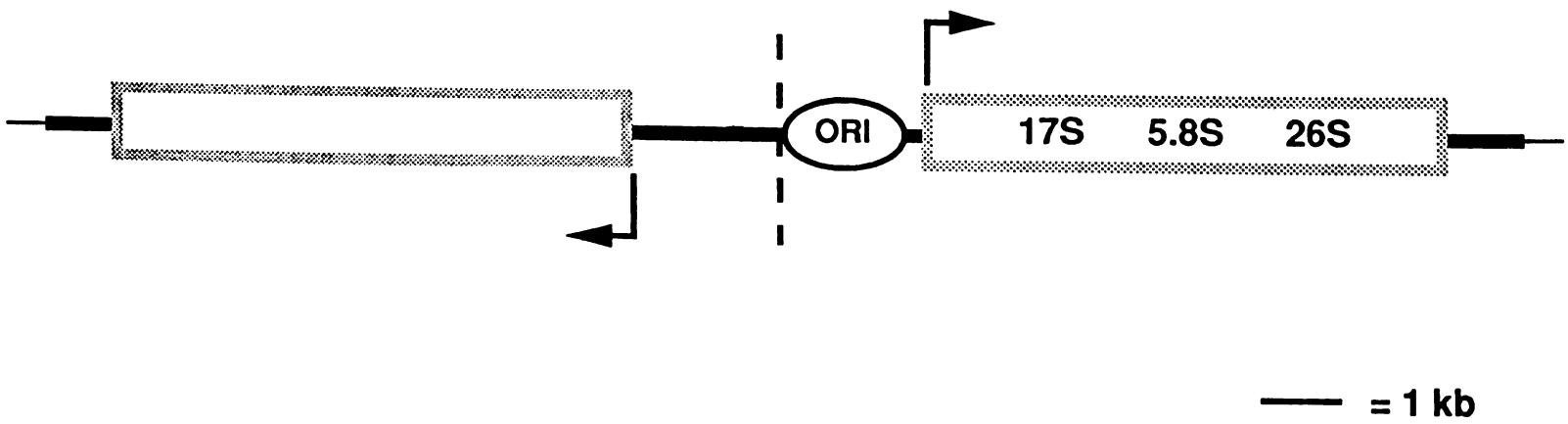
Introduction

The rDNA minichromosome of the ciliated protozoan *Tetrahymena thermophila* has been a useful model system for studies of chromosome maintenance. Single-celled ciliated protozoa have two types of nuclei, a transcriptionally silent germ-line micronucleus and the transcriptionally active macronucleus. The macronucleus is formed from a copy of the micronucleus during conjugation in a process that involves chromosome fragmentation and DNA rearrangement. The rRNA genes are single copy in the diploid micronucleus and are excised from chromosome 2 and formed into a 21 kilobase (kb) palindromic molecule containing two copies of the rRNA genes and flanking DNA sequences in the polyploid (200 chromosomes) macronucleus (Figure 1). This chromosome is amplified to ten thousand copies in the developing macronucleus, and is maintained at this copy number during subsequent vegetative maintenance. The ploidy of the other macronuclear chromosomes is

Figure 1: Structure of the *Tetrahymena* Macronuclear rDNA Minichromosome

This 21 kb molecule is a palindrome, the dashed line marks its center. In most cases the two halves of the palindrome are identical, except for a 28 base pair non-palindromic region at the center of the molecule. The important structural features of the molecule are illustrated for the right half of the palindrome, and are also found on the left half. The heavy black lines indicate the 5' non-transcribed spacers at the center of the molecule, and the 3' non-trancribed spacer at the ends of the molecule. The thin black line represents telomeric DNA. The bent arrow represents the start site of transcription of the rRNA genes. There is a single Pol I transcription unit per half palindrome, its extent is indicated by the open rectangle. The 35 S transcript produced from this region is processed into the 17S, 5.8S and 26S rRNA genes. The order of these within the transcription unit, but not the precise location, is indicated within the rectangle on the right. Finally, each half of the palindrome contains an origin of DNA replication (ORI) within the 5' NTS, indicated by the open oval for the right half of the palindrome. Since each half of the palindrome is identical we can consider the origin region of just one half of the palindrome, see Figure 2.

Structure of the Tetrahymena Macronuclear rDNA Minichromosome



about fifty. These chromosomes are also much larger, making the rDNA minichromosome the smallest and most abundant of the macronuclear chromosomes. It comprises 2% of the total DNA in the cell. In addition, while most other organisms have tandem arrays of rRNA genes, making genetic analysis of an individual locus difficult, the *Tetrahymena* rRNA genes are amenable to genetic analysis because they are single copy in the micronucleus (Gall, 1986; Gorovsky, 1986).

The vegetative replication origin of this molecule has been mapped by electron microscopy and neutral-neutral 2-dimensional gel electrophoresis to the 5' non-transcribed spacer (NTS) of the rRNA genes (Cech and Brehm, 1981; Truett and Gall, 1977) (Kapler and Blackburn, unpublished results) (Figure 1). We have used a genetic screen, based on the existence of a naturally occurring chromosome maintenance mutant carrying a drug resistance marker, to identify other mutants in chromosome maintenance (Larson et al., 1986; Yaeger et al., 1989). We have identified base changes within the 5' NTS in these rDNA maintenance mutants (rmm) which correlate with the maintenance defect (Larson et al., 1986; Yaeger et al., 1989) (this thesis). These base changes are found both in the promoter region (*rmm3* and *rmm8*), and in a region 700 base pairs upstream of the promoter (*rmm1,4,7*), which is within the mapped physical and functional origins (Kapler and Blackburn, unpublished results; Meng-Chao Yao, personal communication). Interestingly, a sequence identical to that of the Type I element mutated in the origin region is also found at the promoter, and this promoter distal

Type I repeat is mutated in *rmm3*. The base changes found in these repeats are either additions (*rmm7*) or deletions (*rmm 1,4,3*) of A residues in a central stretch of 11 As. Figure 2 illustrates the location of the Type I repeats, their sequence and the base changes of the rmm mutants (Challoner et al., 1985; Larson et al., 1986; Yaeger et al., 1989). Repeats 1 and 2 indicate a 400 base pair (bp) tandem duplication found within the rDNA of *Tetrahymena thermophila* that is not found in other Tetrahymenid ciliates. These two regions are greater than 90% homologous with respect to DNA sequence, and DNase I, DMS and KMnO4 footprinting of nuclei indicates that their chromatin structures are virtually identical (Chapter two, this thesis). The physical and functional replication origins localize to a region that includes both repeats (Kapler and Blackburn, unpublished results; Meng-Chao Yao, personal communication). Whether or not they have the same function is not known. To date no maintenance mutants with base changes in the Repeat 1 region have been identified. The presence of the Type I repeats in the promoter region and the origin region, and the genetic evidence from the maintenance mutants that a promoter region Type I repeat and an origin region Type I repeat are important for wild-type chromosome maintenance has suggested a link between rDNA minichromosome replication and rRNA transcription (Larson et al., 1986; Yaeger et al., 1989) (this work). Here we address replication function, transcription function, and protein-DNA interactions at the origin region and the promoter region in *rmm3,* a maintenance mutant which has a deletion of an A within the promoter distal Type I repeat (Figure 2).

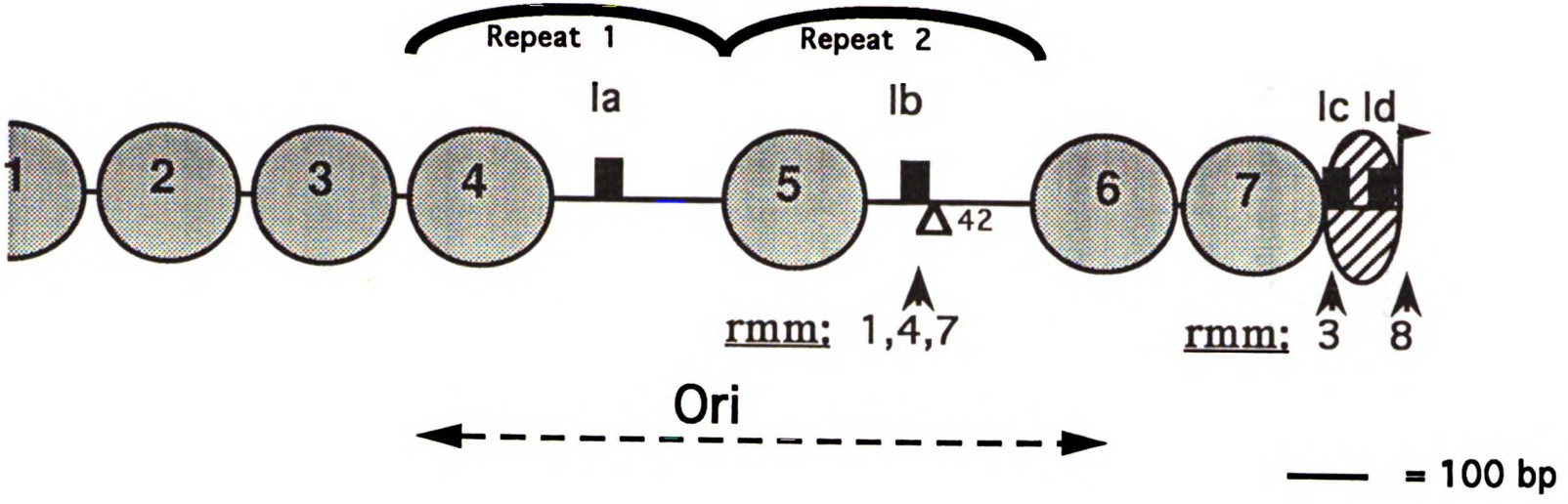
Figure 2: rDNA Maintenance Mutants have Critical Base Changes Withing the 5' NTS

The 1.9 kb 5' NTS of the rDNA minichromosome contains base changes which are associated with a maintenance mutant phenotype. The rough positions of these are illustrated above, the specific location and the specific mutations are given below.

The mutations are located in two regions, a non-nucleosomal region between nucleosomes 5 and 6, and the promoter region.

B strain *Tetrahymena* contain an rDNA allele which is a naturally occurring maintenance mutant, this has a deletion of 42 base pairs between the Type Ic repeat and the adjacent topoisomerase I site (not indicated here). The filled circles indicate the sites of highly positioned nucleosomes within the 5' NTS. The hatched oval indicates the site of an extended footprint at the promoter region.

rDNA Maintenance Mutants have Critical Base Changes Within the 5' NTS



Ia, Ib, Ic = TTTTTTTGGCAAAAAAAAAAACAAAAATAGTAA
Id = ATTCTTTGGCAAAAAAAATAAAAATAATATCAG

rmm1, *rmm4* = - A in Type Ib (-720)
rmm3 = - A in Type Ic (-100)
rmm7 = + A in Type Ib (-720)
rmm8 = G to A 2 n.t. downstream of Type Id (-19)
rmm5, *rmm9* = ? They are not in the 5'NTS
B rDNA has a 42 n.t. deletion adjacent to Type Ib (-700)

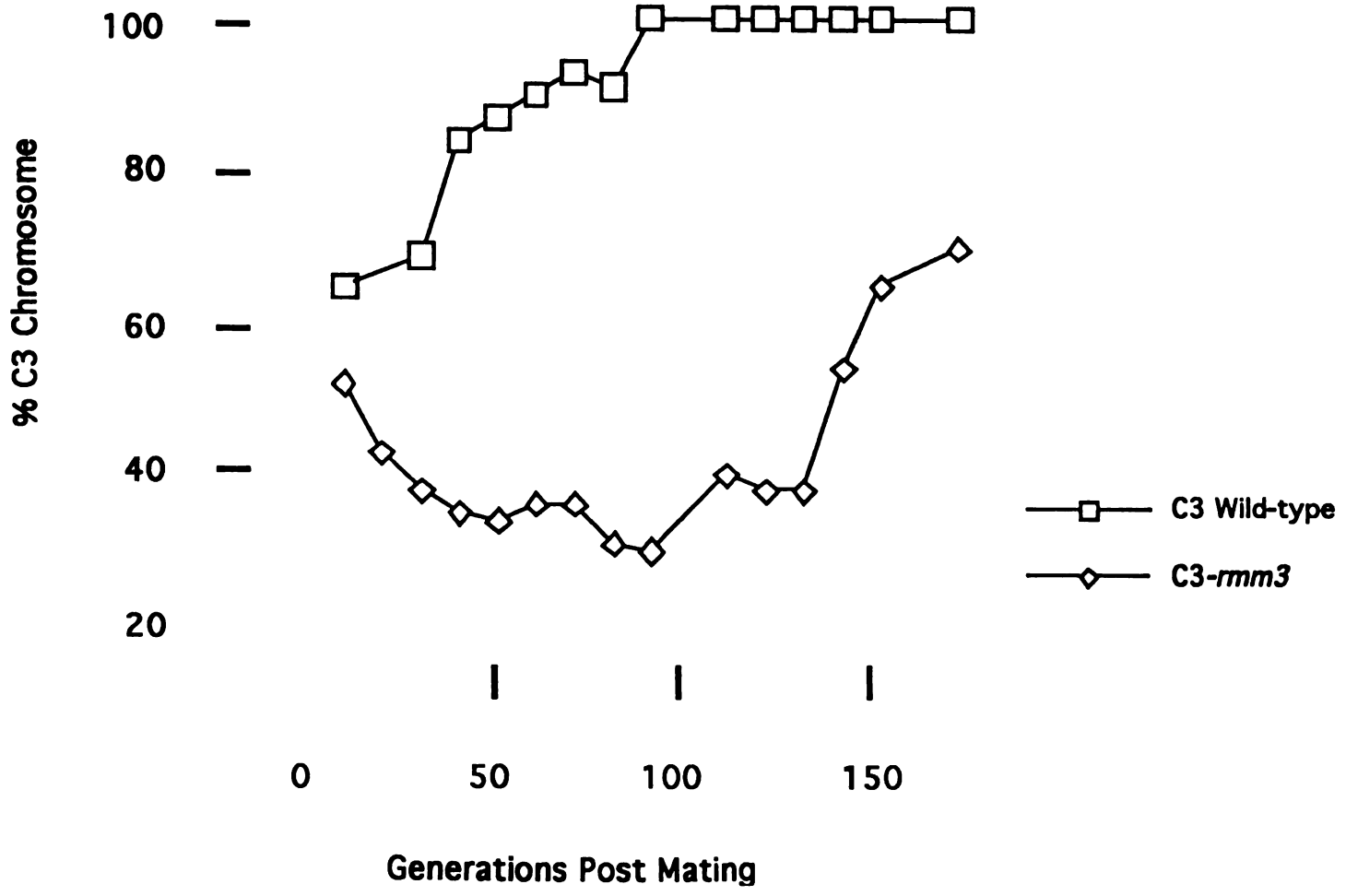
Macronuclear divisions are amitotic and are frequently unequal. As a result an entire allele can be lost from a vegetatively dividing macronucleus over time (Larson et al., 1991; Prescott, 1994). This demonstrates that the macronucleus lacks a mechanism to insure faithful transmission and segregation of genetic information. A copy number control mechanism must exist to maintain macronuclear chromosomes at their ploidy during vegetative growth; however, how this might function has not been elucidated. The unequal division of the macronucleus and the observation of copy number control suggest that chromosomal replication is not strictly confined to once per cell cycle. The loss of the B rDNA allele from the macronucleus of a C3/B rDNA micronuclear heterozygote within 100 generation after mating (Figure 3) further suggests this. Our working hypothesis is that the C3 allele has a replication advantage over the B rDNA allele. However, it has not been shown directly that rDNA replication and not rDNA segregation is responsible for the maintenance phenotype. Density shift experiments do not show evidence of greater than once per cell cycle replication of the rDNA minichromosome in homozygous B strain cells (Truett and Gall, 1977). Our analysis of the *rmm3* maintenance mutant provides the first demonstration of a replication phenotype in a maintenance mutant, and supports our hypothesis that the maintenance phenotype is due to a primary effect on replication.

The unusual features of the *Tetrahymena* macronuclear rDNA minichromosome are the following: it is found in a nucleus which is

Figure 3: An *In Vivo* Competition Assay Demonstrates the Maintenance Defect of C3-*rmm3* rDNA

The wild-type C3 strain C3 491-1a was crossed with the naturally occurring rDNA maintenance mutant B strain SB 1915. Progeny were selected and cells were maintained in log phase growth, DNA was isolated every ten generations. The C3 and B alleles can be distinguished by several restriction fragment length polymorphisms. DNA was cut with Bam HI and the Southern was probed with a 5' NTS probe. The relative intensities of C3 and B specific bands were quantified on the PhosphorImager. In the graph this result is expressed as the percentage of C3 rDNA over time.

An In Vivo Competition Assay Demonstrates the Maintenance Defect of C3-*rmm3* rDNA



amitotic, its mechanism of segregation has not been defined, it is maintained at a copy number of ten thousand molecules per cell by an unknown mechanism, and it may escape cell cycle control at a low level (Larson et al., 1986; Larson et al., 1991; Prescott, 1994). It is likely that one or more of these features have allowed us to observe the effects on chromosome maintenance of the base changes we describe here. Despite the unusual features of the rDNA minichromosome our results are likely to be generalizable to chromosomes of more conventional nuclei, and indicate that long range interactions between a promoter region and an origin region are important for wild-type chromosome maintenance.

Results

Identification of base changes in the 5' NTS of two new maintenance mutants:

We have reported previously the results of a genetic screen to identify mutants in chromosome maintenance, and more recently have repeated the screen in a search for new mutants (Kapler and Blackburn, 1994; Kapler et al., 1994; Larson et al., 1986; Yaeger et al., 1989). The screen is an *in vivo* competition assay between two naturally occurring rDNA alleles, C3 and B. The B allele carries a drug resistance marker. Initially, a cell heterozygous for the two alleles contains roughly equal amounts of each in the macronucleus (Orias and Bradshaw, 1992). However, the B rDNA allele is lost from the cell within seventy to one hundred generations after mating, even in the presence of a selection for the chromosome. We generated new mutants in rDNA maintenance by mutagenizing C3

strain cells and crossing them with B strain cells. C3 rDNA maintenance mutants amplify the mutant C3 allele but do not maintain it as well as wild-type C3 rDNA in the presence of B rDNA. These are identified by crossing mutagenized C3 strain cells with B strain cells and scoring progeny for drug resistance after 70 generations. Drug resistant progeny still retain the B allele and thus have a mutant C3 allele. Four new rDNA maintenance mutants (*rmm 5,7,8,9*) were generated (Kapler and Blackburn, 1994; Kapler et al., 1994). We sequenced the entire 5' NTS of the four new rDNA maintenance mutants. No base changes from wild-type C3 rDNA were found in *rmm5* or *9* , indicating that sequences other than those in the 5' NTS are important for chromosome maintenance. In *rmm7* an addition of an A in a run of eleven A's was found in the Type Ic repeat (Figure 2). In *rmm8*, a G to A transition was found in a run of 6 G's that begins with the last G of the Type Id repeat at -21 from the start site of transcription and extends toward this site. This is within a larger phylogenetically conserved sequence that extends from -135 to +35 (Challoner et al., 1985). Thus, these two new maintenance mutants had base changes within the replication origin region and the rRNA gene promoter.

The *rmm7* base change is found in the same element that was altered in the previously characterized *rmm1* and *rmm4* mutants (Figure 2). The *rmm8* mutant has a base change in the promoter region, which was also mutated in the previously isolated *rmm3* mutant (Larson and Orias, personal communication) (Figure 2). The identification of base changes in a Type I repeat and in a promoter

region in two new independent maintenance mutants strongly supports the hypothesis that these two elements are important for wild-type rDNA minichromosome maintenance. In order to determine the role of the Type I repeat and the promoter region in chromosome maintenance we analyzed rRNA transcription, rDNA replication, and protein-DNA interactions at the origin region and the promoter region in the *rmm3* mutant, which has a base change in the Type Ic repeat in the promoter region (Figure 2).

The *rmm3* rRNA promoter region mutant has a chromosome maintenance phenotype:

We confirmed the effect of the *rmm3* mutation on rDNA maintenance in the presence of the B rDNA allele in a backcrossed homozygous *rmm3* strain. Wild-type C3 or C3-*rmm3* strain cells were crossed with B strain cells, progeny were maintained in log phase growth and DNA was isolated every 10 generations. B strain cells have an SphI site in the 5' NTS that C3 cells lack (Engberg and Nielsen, 1990). DNA was cut with SphI, analyzed by Southern, and the relative intenisites of the bands used to determine change in the ratio of rDNA alleles over time was quantitated on the phosphorImager. The data for both crosses is plotted as the per cent of C3 chromosome in the macronucleus over time. By 100 generations C3 rDNA out-competes all the B rDNA in the cell. However, as expected, C3 -*rmm3* rDNA does not out-compete the B rDNA; instead, both alleles persist until recombination generates a C3 chromosome that has lost its maintenance defect (Figure 3). This homozygous backcrossed strain, *rmm3* Rd 2 #1 was used for the

experiments reported in this work. In order to determine whether the deletion of an A at the Type Ic repeat was responsible for the maintenance defect of the C3-*rmm 3* chromosome, PCR primers specific for the C3 promoter region were used to amplify DNA from both crosses over time. Sequencing of the PCR products revealed that the Type Ic repeat of the C3 chromosome in the C3-*rmm3* cross was restored to wild-type in parallel with the reversion of the maintenance defect, supporting the conclusion that the base change at Type Ic is responsible for the maintenance phenotype of these cells (data not shown).

The *rmm3* rRNA promoter region maintenance mutant has a replication phenotype:

Neutral-neutral two-dimensional (2-D) gel analysis has been used to analyze replicating intermediates of vegetative cells and supports the conclusions from electron microscopy and BrdU labelling that the vegetative origin of the rDNA minichromosome is located within the 5' NTS of the rRNA genes (Kapler and Blackburn, unpublished results). A distinctive feature of the observed pattern are replication fork pauses within the 5' NTS. Multiple pauses are not found in the 2-D gel patterns of the rDNA of organisms that also have localized initiation of DNA replication, pea and *Physarum* (Benard et al., 1995; Hernandez et al., 1993). Multiple replication fork pauses are seen in the patterns of rDNA replicating intermediates in human and *Xenopus* rDNA. However, these pauses have been attributed to replication fork collisions. In these organisms, with much larger spacer regions, the initiation of DNA

replication occurs at multiple positions in a single 5' NTS (Hyrien et al., 1995; Little et al., 1993). 2-D gels indicate that there is a single initiation site in *Tetrahymena* (Kapler and Blackburn, unpublished results), so the pauses are unlikely to be due to colliding forks. The cause of the replication fork pauses is not known, their presence may make more precise mapping of the replication origin difficult.

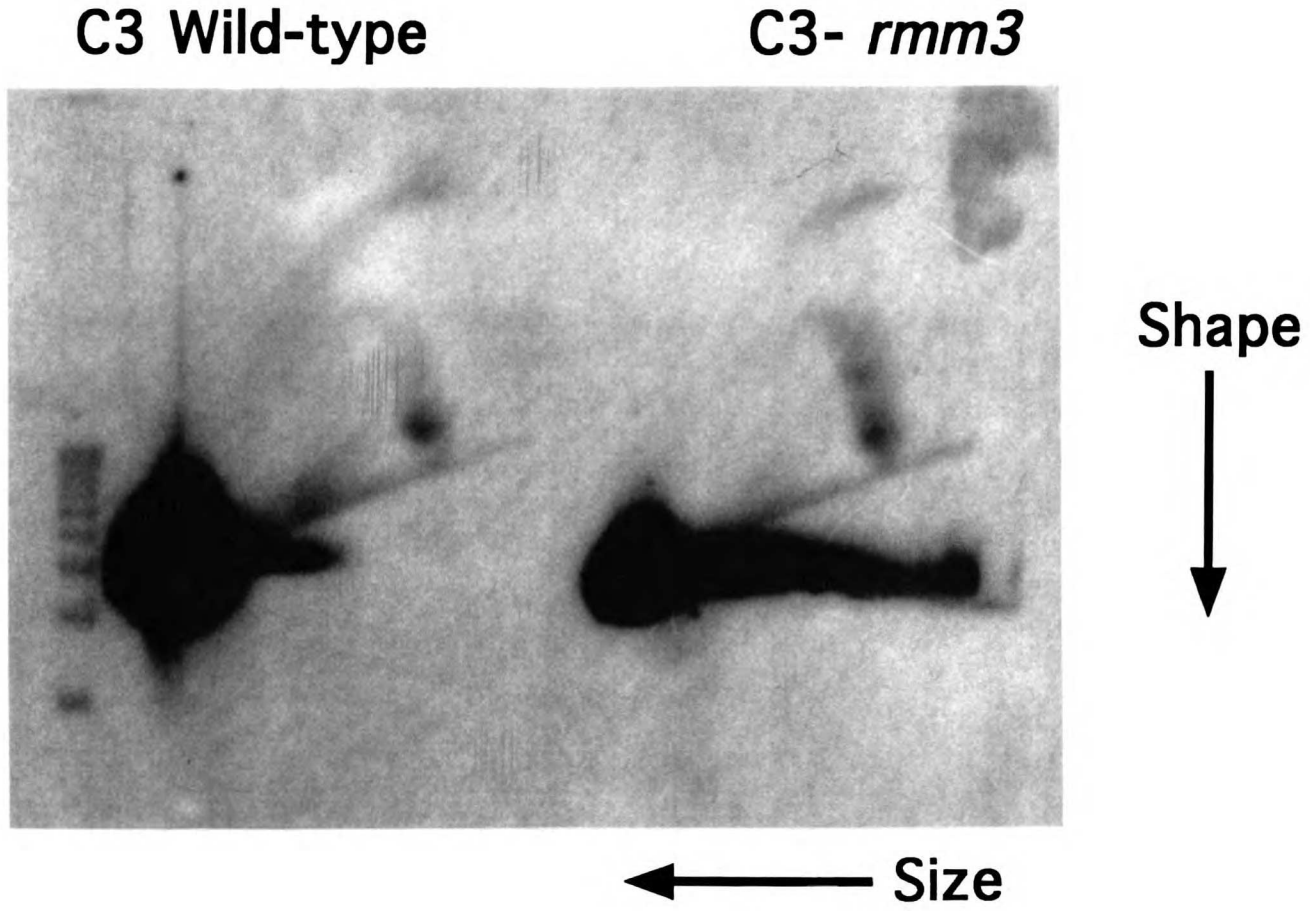
I have used neutral-neutral 2-D gels to analyze replicating intermediates of C3 wild-type and C3-*rmm3* homozygous cells. Cutting with Hind III generates a palindromic 4.2 kb fragment that spans the 5' NTS regions of both halves of the rDNA palindrome (Figure 1). In neutral-neutral 2-D gels DNA is run in the first dimension under conditions that separate molecules by size, and in the second dimension under conditions that separate molecules of the same size by shape. Molecules containing replication forks or bubbles are retarded in the second dimension and run above the arc of linear double stranded DNA. Bubble-containing fragments run in a high arc, and fork-containing fragments run in a low arc that returns to the arc of linears.

2-D gels probed for the central Hind III fragment of the rDNA minichromosome demonstrate a bubble to Y (fork)-arc pattern, indicating that a replication bubble is located asymmetrically within the fragment (Figure 4). Comparison of 2-D gels of C3 wild-type and C3-*rmm3* log phase cells indicates that replicating intermediates accumulate at one position in the 5' NTS to a greater

Figure 4: 2-D Gels Indicate Enhanced Accumulation of C3-*rmm3* rDNA Replicating Intermediates in 5' NTS

Neutral-neutral 2-D gels were run as described in Materials and Methods. DNA from log phase cells was restricted with Hind III which generates a 4.2 kb fragment that spans both 5'NTS s within the palindrome. The resulting pattern of replicating intermediates is that seen when an origin of replication is located asymmetrically within a restriction fragment and is termed a bubble to Y transition. The accumulation of replicating intermediates at a pause site results in a spot along the fork arc, see text for description.

2-D Gels Indicate Enhanced Accumulation of C3- *rmm3* rDNA Replicating Intermediates in 5' NTS



extent in C3-*rmm3* rDNA than in wild-type cells (Figure 4). In each panel the very dark spot is the unreplicated DNA of the 4.2 kb central Hind III fragment. The thin line of non-specific hybridization extending slightly upward and to the right is the arc produced by the bulk, linear, unreplicated DNA. The high arc is the bubble arc, and the low arc which returns to the arc of linears is the fork arc. Note the spot at the end of the fork arc in both panels, adjacent to the arc of linears. This indicates that there is an accumulation of replicating intermediates due to a replication fork pause near the end of the restriction fragment. In C3-*rmm3* cells there is an additional hybridizing spot above this indicating another replication fork pause within the 5' NTS. On a darker exposure this can also be seen in the wild-type cells, but it is more intense in the mutant cells. Thus, C3-*rmm3* cells demonstrate an enhanced accumulation of replicating intermediates at this position, indicating an accentuated replication fork pause. This is the first demonstration of a replication phenotype in an rmm mutant. The absence of another origin on the molecule that could serve as a control for the initiation of replication has the result that we can make no conclusion about the relative initiation frequencies in C3 wild-type and C3-*rmm3* strain cells. We can only compare 2-D gels for replication elongation.

The *rmm3* rRNA promoter region mutant has no defect in the initiation of transcription:

The identification of base changes in chromosome maintenance mutants at the promoter region (*rmm3* and *rmm8*), and specifically in a conserved sequence element found both at the promoter and at a candidate origin region (in the Type Ib- origin region element in *rmm1,4,7* and in the Type Ic-promoter region element in *rmm3*, see Figure 2) suggested a link between replication and transcription. The data from the 2-D gels presented above indicates that there is a replication defect in the *rmm3* mutant. We also analyzed transcription in C3-*rmm3* homozygous mutant cells. As mentioned above, the *rmm3* base change responsible for the maintenance phenotype is similar to that of the *rmm1,4* and *7* mutants in that it is a deletion or an addition of an A residue in a Type I repeat. However, the *rmm1, 4* and *7* base changes are in an upstream Type I repeat at -700 base pairs from the start site of transcription. No role for these repeats in transcription has been demonstrated. We thought it more likely that the promoter distal Type I repeat, Type Ic, located from -84 to -116 from the start site of transcription would have direct role in transcription. Analysis of RNA Polymerase I (Pol I) transcription in *Tetrahymena* is limited and indicates an essential role *in vitro* for only the downstream half of the promoter proximal Type Id repeat (Higashinakagawa, personal communication). However, human, mouse, and *Xenopus* Pol I promoters have a bipartite structure that consists of an essential core element near the start site of transcription and an upstream element located about 100 base pairs away which enhances transcription (Heix and

Grummt, 1995). If this structure is conserved in *Tetrahymena*, the Type Ic repeat could enhance Pol I transcription. Thus, it seemed most likely that of the four *rmm* mutants with a base change at a Type I repeat, *rmm3* cells would exhibit a detectable difference in transcription. We examined transcription in two ways. We performed run-ons in wild-type and maintenance mutant cells, and we analyzed a previously described, small, distinct rRNA transcript that normally accumulates in cells and is within the external transcribed spacer at the 5' end of the transcription unit (Figure 5a,b,c).

I assayed the initiation of transcription in cell ghosts prepared from log phase or from starved cells. Run-on assays were performed according to the procedure of Love et. al. (Love et al., 1988). Cell ghosts were incubated with rNTPs and ^{32}P-UTP for 2 minutes at 30°C. Labelled RNA was treated with DNase I and Proteinase K, and cleaved by mild NaOH treatment to prevent labelled RNA molecules from hybridizing to more than one probe. The RNA was then hybridized to filters containing PCR generated DNA probes spanning the 35S RNA Pol I transcript. The locations of the PCR generated DNA probes used are indicated in Figure 5a. The 35S transcript is processed to produce the 17S, 5.8S, and 26S rRNAs (Engberg et al., 1984; Kister et al., 1983). It is believed to be the only Pol I transcript; however, the IF probe corresponds to a region which has a distinct transcript(s) on Northern analysis (Kister et al., 1983). This transcript(s) has been described previously and has now been characterized more fully by us (see below). It has been termed the

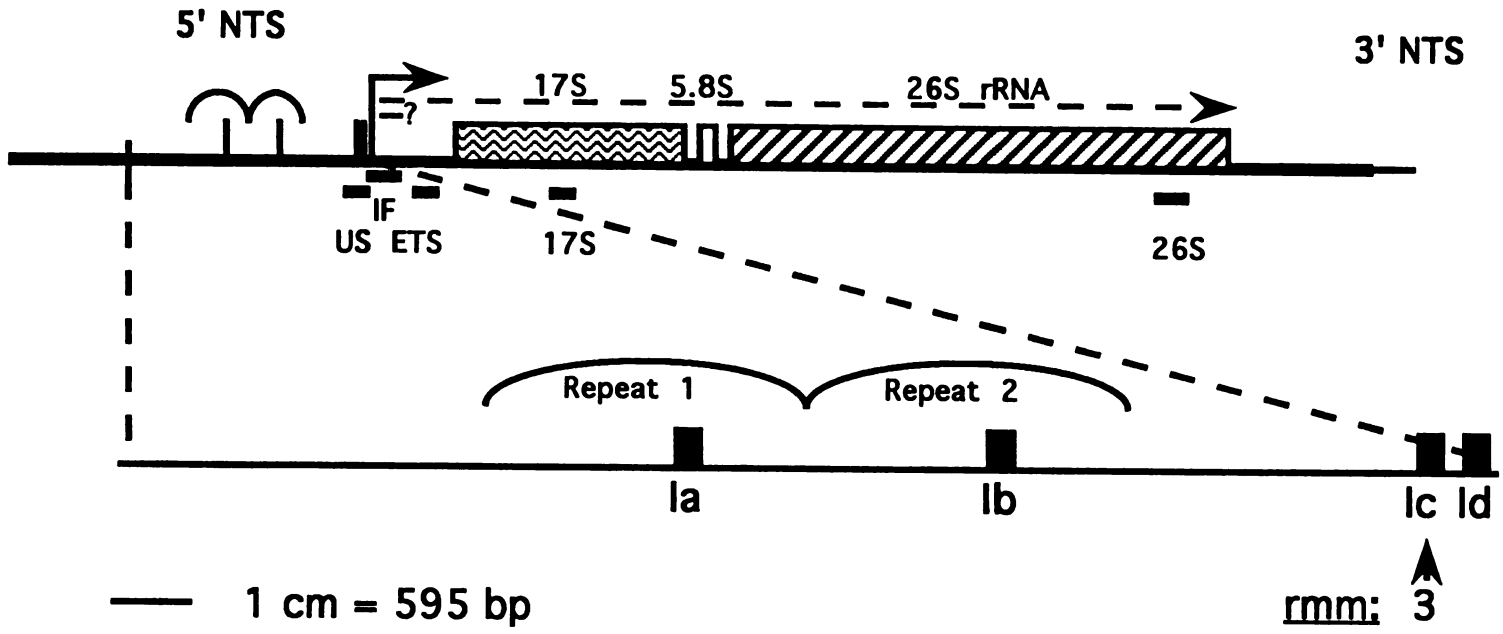
initiator fragment because it does not appear to be a processing product of the 35S transcript. It has been suggested that it is an attenuation product of the Pol I promoter (Kister et al., 1983); similar small Pol I transcripts that do not appear to be processing products of the 35S RNA are found in human cells (Reichel and Benecke, 1984). I included this probe in an attempt to identify attenuation at this Pol I promoter, and to determine if there was any difference in the production of this initiator fragment transcript between wild-type and *rmm3* homozygous cells. US is an upstream region probe that extends from - 222 to - 4 from the start site of transcription. The ETS probe is located within the external transcribed spacer of the 35S transcript, and the 17S and 26S probes are within the coding sequences of those genes. The location of the *rmm3* base change is indicated in the blow-up below the schematic (Figure 5 and Figure 2).

The first part of Figure 5b presents the data for log phase cell ghosts. Blots were scanned on the phosphorimager. For each data point the number of pixels was divided by the number of U residues in the probe, and normalized to the level of 5S gene hybrdization on that blot. As suggested by eye, and as confirmed by the normalized phosphorimager data, there was no significant difference between log phase *rmm3* strain cells and wild-type C3 strain cells in the initiation of transcription, or in the density of polymerases throughout the gene. There is no evidence that the initiator fragment is an attenuation product, since polymerase density does

Figure 5a: rDNA Chromosome Schematic: One Half of the Palindrome

The schematic of one half of the palindrome shows the transcribed region of the rDNA and the position of the probes used to analyze run-on transcription in wild-type and C3-rmm3 maintenance mutant strains. These are US: upstream probe, IF: initiator fragment, ETS: external transcribed spacer, 17S: 17S rRNA, 26S: 26SrRNA. The bent arrowhead indicates the start site of transcription, the long dashed line represents the 35 S primary rRNA transcript which is processed into the 17S, 5.8S and 26S rRNAs. The positions of these genes are indicated by the wavy, open and hatched boxes, respectively. The dash with a question mark beside it represents the "initiator fragment". Two transcripts which hybridize to this region have been identified (see figure 5c), but it is not clear whether they represent bona fide transcripts or are processing products. The 5' NTS and the 3' NTS are indicated. The blow-up of the 5' NTS indicates the position of the rmm3 mutation in the promoter distal Type Ic repeat. The thin line at the end of the 3' NTS represents the telomere.

rDNA Chromosome Schematic: One Half of the Palindrome



not fall off detectably between the regions identified by the IF and ETS probes. The 2X 17S slots were loaded with twice as much DNA as the 17S slots. The fact that hybridization of labelled RNA did not increase significantly indicates that DNA was not limiting on the blot.

The Pol I promoter is extremely active in log phase cells. I considered the possibility that the effect of the *rmm3* base change on transcription is small and would be detectable only in a situation in which the overall level of Pol I transcription was reduced. Pol I transcription is reduced significantly in starved cells (Engberg et al., 1972). Figure 5b also presents the data for run-on transcription in wild-type and C3-*rmm3* starved cell ghosts. As above, the phosphorimager data was normalized both for the number of U residues in the probe and to the level of 5S gene transcription on the same blot. Comparison of the normalized values of transcription of the IF fragment indicates that there is no difference in the initiation of transcription between wild-type and C3-*rmm3* homozygous cells. Comparison of the normalized values of hybridization of the ETS and the 17S fragments to labelled RNA from the two strains suggests that there may be a difference in polymerase density downstream of the IF fragment between wild-type and C3-*rmm3* strain cells. There is greater hybridization to the ETS probe than to the IF probe in wild-type cells. In starved C3-*rmm3* homozygous cells the IF and the ETS probe hybridize to the same extent. Similarly, the 17S probe hybridizes to labelled RNA to a greater extent than the IF probe in wild-type cells, but the probes

Figure 5b: *rmm3* Homozygotes Are Not Defective in the Initiation of Transcription

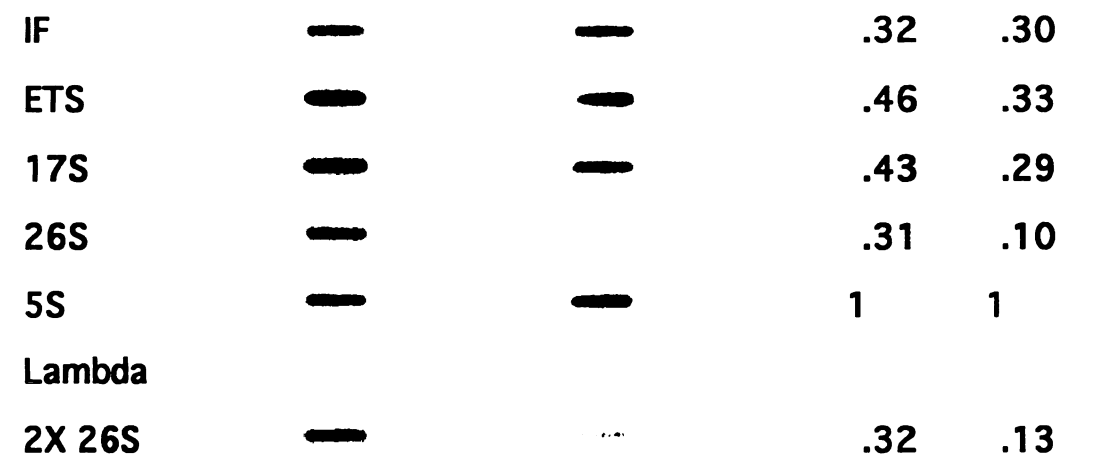
Run-on Assays were performed on cell ghosts in log phase and in starved cells. See Materials and Methods for details. In addition to the rDNA PCR generated probes described above the following DNAs were hybridized to the blot, 5S rRNA gene probe (Pol III transcribed), a gamma tubulin probe (Pol II transcribed), and lambda DNA. In order to ensure that DNA was not limiting on the blot twice as much DNA from one sample was hybridized to one slot. The signal from these slots did not go up significantly indicating that DNA was not limiting on the blot. The run-ons were scanned on the PhosphorImager. The raw data was normalized twice, first to the number of U's in the transcribed region, then to the 5S gene hybridization of that blot.

rmm3 Homozygotes Are Not Defective in the Initiation of rDNA Transcription

Log Phase Cell Run-Ons:

	wt	*rmm3*	# nt Probe	# Us	Normalized PI Data wt	*rmm3*
US			219			
IF			268	71	1.6	1.3
ETS			250	70	1.8	1.8
17S			240	73	1.6	1.6
26S			262	74	3.0	3.5
5S			120	21	1	1
TUB						
Lambda						
2X 17S					2.3	1.5

Starved Cell Run-Ons:



	Normalized PI Data	
	wt	*rmm3*
IF	.32	.30
ETS	.46	.33
17S	.43	.29
26S	.31	.10
5S	1	1
Lambda		
2X 26S	.32	.13

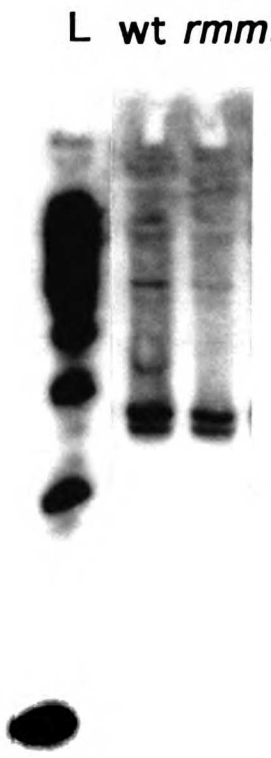
hybridize to the same extent in C3-*rmm3* cells. These values were highly reproducible for the same set of cells. However, the variability of the hybridization observed in these assays (see Figure 5b, compare 17S and 17S 2X) is close to the hybridization difference observed between wild-type and C3-*rmm3* strain cells, so we can come to no firm conclusion about the difference between the two with respect to the production of the 35S transcript. In contrast, there is a clear difference in polymerase density between the two strains in the 26S rRNA gene region at the end of the 35S transcript. The 26S rRNA gene is 3.8 kb and the probe spans the region from 500 to 240 nt from the end of the 26S rRNA gene, the 35S transcript ends 14 nts downstream of this (the 35 S transcript is 6.7 kb). In log phase cells this probe hybridized twice as efficiently to labelled RNA, and there was no significant difference in hybridization between the two strains (Figure 5b). In starved wild-type cells the 26S probe hybridizes to labelled RNA to the same extent as the IF probe. However, in starved C3-*rmm3* strain cells hybridization to this probe is decreased three fold. This was consistent between duplicate samples. The 2X 26S control contains twice as much DNA per slot and shows the same amount of hybridization as the 26S slot in each strain. Thus, DNA is not limiting on the blot, and the reduced hybridization in the 26S lane of *rmm3* is not due to underloading of the DNA probe. These data indicate that polymerase density in the 26S gene region of starved C3-*rmm3* strain cells is significantly reduced in comparison to wild-type C3 cells.

Finally, I analyzed the initiator fragment transcript(s). This small transcript is not made in an in vitro transcription reaction, and it does not appear to be a processing product of the 35S gene (Kister et al., 1988). It was reported as a 230 nucleotide transcript on an agarose gel (Kister et al., 1983). On polyacrylamide Northerns we found two transcripts in this size range. In order to attempt to map these transcripts I sized these on a sequencing gel and estimated their lengths to be 225 and 240 nucleotides, respectively. Next, I sequentially hybridized the Northern to several different probes. Neither transcript hybridized to a probe that extended from - 21 to - 4 from the start site of transcription and both hybridized to an oligo probe that extended from + 40 to + 60. A probe made to distinguish the transcripts extended from position + 219 to + 241 and hybridized only to the longer transcript (data not shown). I did not attempt to determine the 5' end of the transcripts by primer extension, therefore I cannot be certain that they have exactly the same 5' end, or that their 5' ends correspond to +1 of the 35S transcript. However, we believe our probe would have detected transcripts that initiated at -16 and upstream. I conclude that the two transcripts initiate at or in the region of +1 and that they extend for approximately 225 and 240 nucleotides, respectively. Northern analysis of these two transcripts in log phase wild-type and C3-*rmm3* cells revealed no difference in size or abundance of either transcript (Figure 5c). There was no significant difference in the transcripts between the two strains in starved cells (data not shown).

Figure 5c: "Initiator Fragment" Transcripts Are Not Altered in C3-*rmm3* Cells

The only significant transcripts identified with a 5' NTS probe that extends twelve nucleotides into the 35S coding region correspond to the previously described initiator fragment. Here we show that this initiator fragment is really two transcripts, and that these are indistinguishable in wild-type and C3-*rmm3* log phase cells. The Northern was probed with primer 12 which extends from +60 to + 40. L is the 100 base pair ladder labelled with ^{32}P-ATP.

"Initiator Fragment" Transcripts Are Not Altered in C3-*rmm3* Cells



Loss of DMS reactivity of an A residue in the non-nucleosomal Domain 1 and Domain 2 of the *rmm3* origin region:

I have used DNase I, DMS, and KMnO4 to examine protein-DNA interactions in the 5' NTS of nuclei of wild-type and maintenance mutant cells. In chapter two I described the overall chromatin structure of the 5' NTS. Two thirds of it is covered by seven highly positioned nucleosomes. At the region of the promoter there is an extensive DNase I footprint and DMS protections and enhancements that extend from the start site of transcription to - 126. Finally, there are two non-nucleosomal regions which contain DMS and KMnO4 footprints at the three Type III repeats per domain which are the sites of action of Topo I (Pan et al., 1995) this thesis. These latter two non-nucleosomal regions, termed Domain I and Domain 2, lie within Repeat 1 and Repeat 2 respectively, and are good candidates for the sites at which origin recognition factors might bind (Figure 2). 2-D gels and transformation experiments have localized the physical and functional origins of replication to the Repeat 1/Repeat 2 region of the 5' NTS (Kapler and Blackburn, unpublished results, Meng-Chao Yao, personal communication). Apart from the DMS and KMnO4 footprints at the Type III repeats the only clear footprint in the Domain 1 and Domain 2 regions is a DMS reactive A residue located sixty-one nucleotides upstream of the 5' boundary of the Type I repeat in each domain (Figure 6a, see below).

Nuclei (+ lanes) or naked DNA (- lanes) were incubated in 10 mM DMS for the indicated times at 25°C. Methylated DNA was cleaved and primer extended in the region of the Type Ib repeat. Sequencing

of plasmid DNA (A T lanes) allows determination of the exact nucleotide positions of the cleavages. rDNA nucleotide numbers are indicated on the side, with primer extension copying the coding strand of the rRNA genes and proceeding toward the center of the molecule. The Type Ib repeat is just off the gel. The brackets indicate the approximate extent of the positioned nucleosome in the Repeat 2 region, nucleosome 5. Domain 2 contains the Type Ib repeat, is located downstream of nucleosome 5 and extends to nucleosome 6, and is the non-nucleosomal region of Repeat 2 (Figure 2). The upstream borders of these regions are labelled on the side of the figure above the corresponding downward facing arrow. DMS methylates the N-7 position of G residues, which is found in the major groove of double-stranded DNA. Both naked DNA and chromatin have a G cleavage ladder. Vertical arrows indicate nucleotides that have reacted with DMS in chromatin to a greater extent than in naked DNA, these are identified by comparing the - and + lanes of each time point. These are the DMS footprinted residues and they are found in three different sites. High in the gel are the cluster of previously described Type III repeats of the Domain 1 region. Analogous sites are found in Domain 2. These are the sites of action of topoisomerase I (Bonven et al., 1985; Challoner et al., 1985). Each Type III repeat is footprinted at a single G, two nucleotides downstream of the site of Topo I cleavage (this thesis). The rDNA nucleotide number of each cleaved residue is indicated adjacent to the corresponding arrow. DMS can also methylate the N-3 position of A s in double-stranded DNA, which falls in the minor groove, and the N-1 position of single stranded As, which is base paired in

Figure 6a: C3 Wild-Type rDNA Has a DMS Reactive A at Nucleotide 1132 in Domain 2

Wild-type C3 strain nuclei were footprinted with 10 mM DMS. Methylated DNA was cleaved with heat or pyrrolidine and was analyzed by primer extension.

- indicates naked DNA

+ indicates chromatin

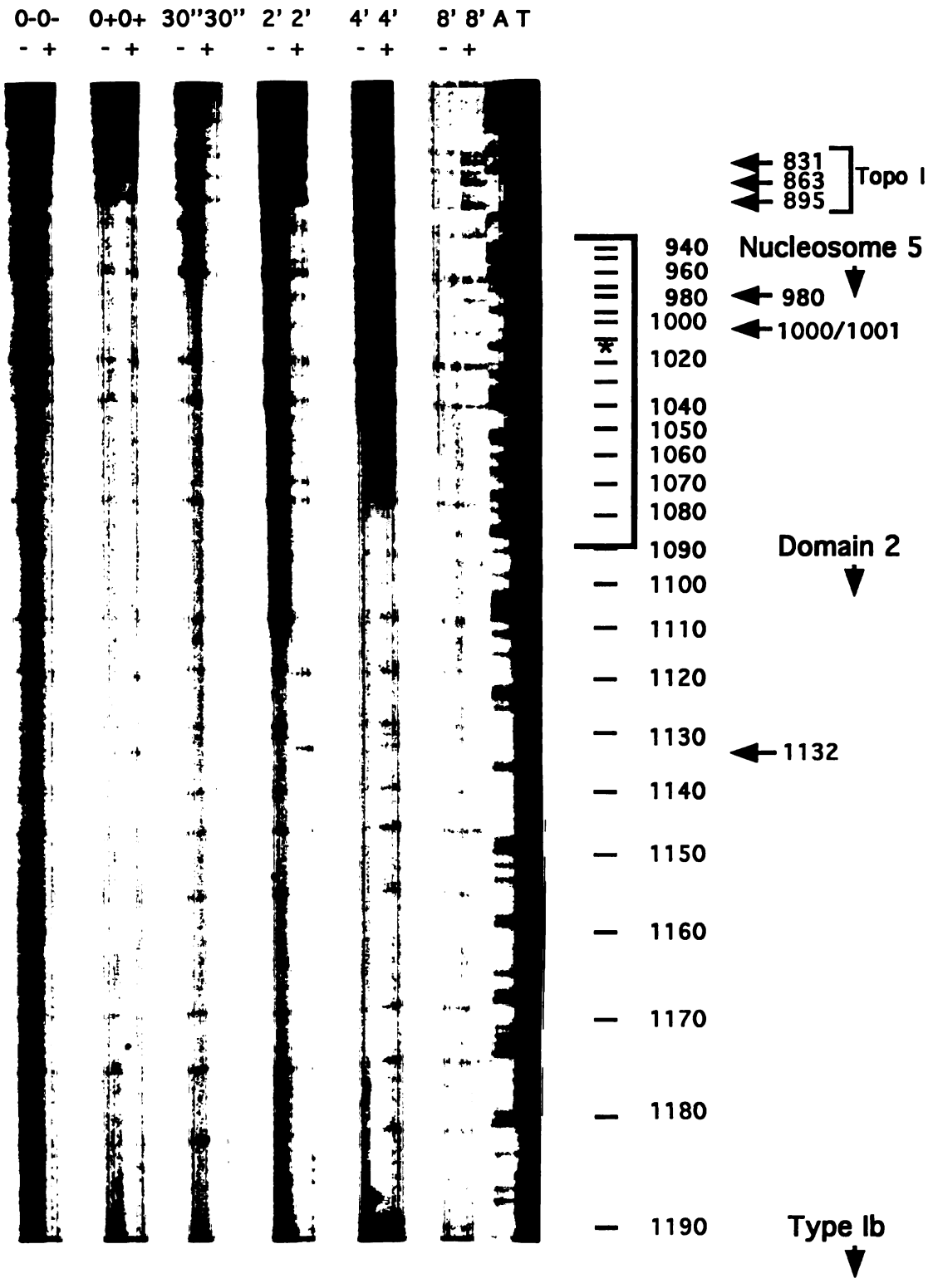
0- indicates samples that were not treated with DMS and were quenched at 0 time.

0+ indicates samples to which DMS was added after samples were quenched at 0 time.

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

In Figure 6a the footprinted region is within Repeat 2 and extends from just 5' of the Type Ib repeat towards the center of the molecule, through nucleosome 5. The particular region is indicated on the far right, the position of nucleosome 5 is bracketed. Actual nucleotide numbers appear to the side.

C3 Wild-Type rDNA Has a DMS Reactive A at Nucleotide 1132 in Domain 2



double-stranded DNA. Within the nucleosome 5 region there are two DMS footprinted A residues. These DMS reactive A residues were described in chapter two. It is likely that these are methylated at the N-3 position, because the corresponding T residues on the opposite strand is not footprinted with $KMnO_4$, which detects single-stranded Ts. Finally, within Domain 2 a single A residue is footprinted at nucleotide 1132. Two G residues, equally reactive in naked DNA and chromatin, are seen at nucleotides 1128 and 1130. At 1132 the A is reactive in chromatin and not in naked DNA (compare - and + lanes throughout the time course). It is possible that this nucleotide is methylated at the N-1 position, because the T on the opposite strand is reactive to KMnO4 (data not shown). Repeats 1 and 2 are highly homologous in sequence, and without exception residues footprinted in one repeat are also footprinted in the other. Thus, the corresponding A residue in Domain I of Repeat 1 is also reactive to DMS, and the corresponding T is reactive to $KMnO_4$ (nucleotide 701, data not shown). In contrast to the results for wild-type cells, these A residues are not footprinted in log phase homozygous C3-*rmm3* strain cells (Figure 6b). This is shown for the DMS reactive A of Domain 2. Comparison of the footprints of the Topo I sites and nucleosome 5 in wild-type and *rmm3* strain cells reveals that a clear footprint is seen at these sites in each cell type. In contrast, nucleotide 1132 is footprinted in wild-type cells and not in *rmm3* strain cells. Similarly, the corresponding nucleotide in Domain 1, nucleotide 701, is not footprinted by DMS in C3-*rmm3* strain cells, and is footprinted in C3 wild-type cells. We have sequenced the 5' NTS of the *rmm3* homozygous cells and

confirmed that nucleotides 1132 and 701 are A residues on the coding strand (data not shown). This is the only strain in which we have seen a loss of the 1132 footprint. We have footprinted other maintenance mutants (*rmm1,* and *8*), and cells transformed with rDNA vectors in which recombination has generated tandem arrays of 5' NTS sequences. These tandem arrays consist of C3 wild-type rDNA sequence with the exception of a +G mutation in the run of 6 Gs from - 16 to - 21 from the start site of transcription which abolishes promoter function (Pan and Blackburn, 1995; Pan et al., 1995). Thus, the loss of the DMS footprint at 1132 is specific to *rmm3* strain cells.

Loss of the Wild-Type DMS Promoter Footprint in *rmm3* Strain Cells:

DMS treated *rmm3* chromatin DNA and naked DNA were also primer extended at the promoter region. As mentioned above, the promoter has not been clearly delineated by functional assays in *Tetrahymena*. We have described previously an extensive DNase I footprint of the region that extends from the start site of transcription to -120 (Pan et al., 1995). There are protections and enhancements in this region that encompass the promoter Type Ic and Type Id repeats (see Figure 8a). Since the *rmm3* strain cells have a deletion of an A in a stretch of 11 As in the promoter distal Type Ic repeat, we investigated the effect of this mutation on the promoter footprint.

The wild-type DMS promoter footprint on the coding strand of the rRNA genes is seen in Figure 7a. DMS treated DNA has been primer

Figure 6b: C3-*rmm3* rDNA Has No DMS Reactive A at Nucleotide 1132 in Domain 2

C3-*rmm3* homozygous cell nuclei were footprinted with 10 mM DMS.

- indicates that naked DNA was reacted with DMS prior to cleavage and primer extension

+ indicates that nuclei were incubated with DMS prior to cleavage and primer extension.

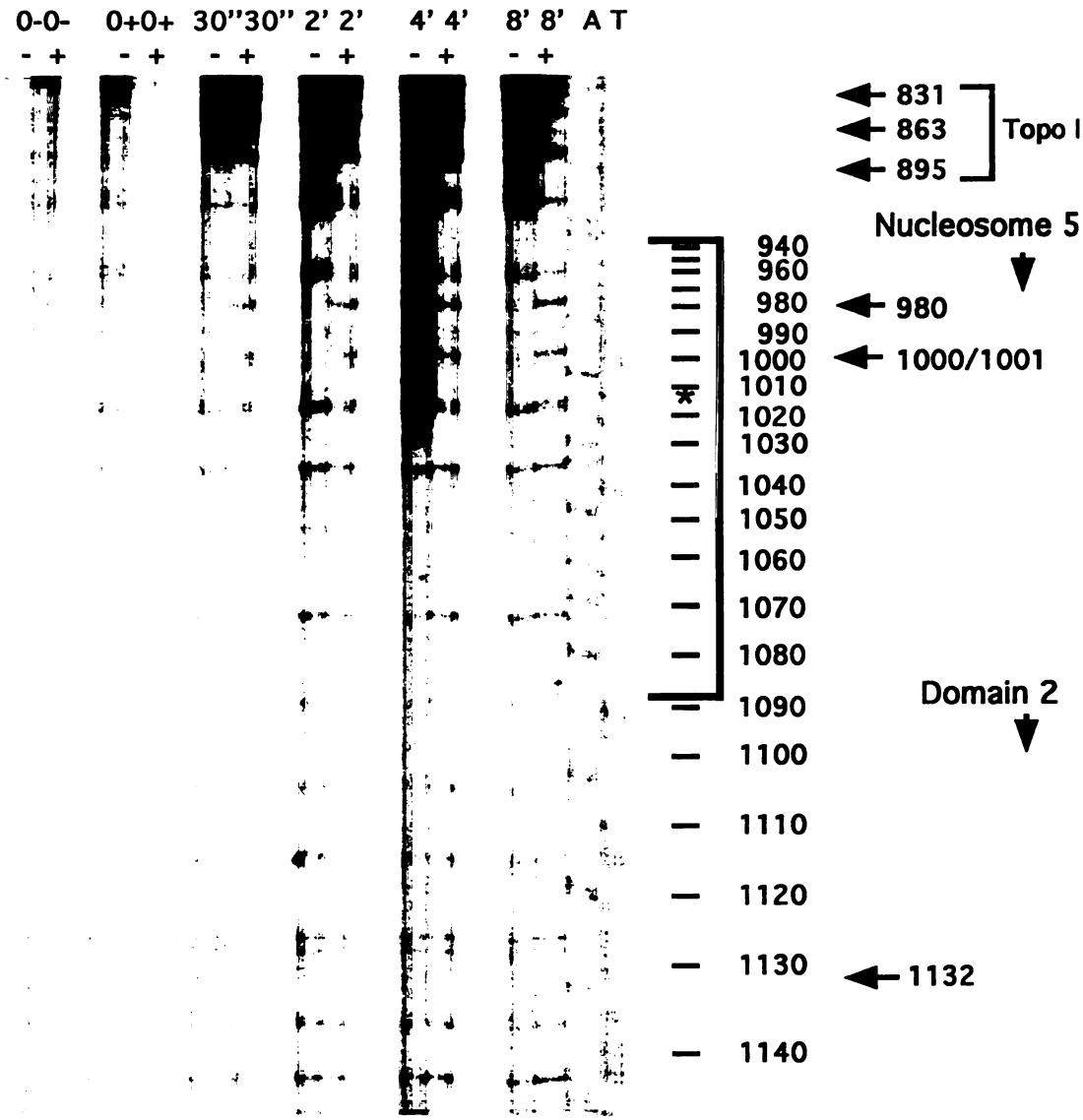
0- indicates samples that were not treated with DMS and were quenched at 0 time.

0+ indicates samples to which DMS was added after samples were quenched at 0 time.

Other samples were treated for 30 seconds, 2 minutes, 4 minutes or 8 minutes, as indicated.

In Figure 6b the footprinted region is within Repeat 2 and extends towards the center of the molecule, through nucleosome 5. The particular region is indicated on the far right, the position of nucleosome 5 is bracketed. Actual nucleotide numbers appear to the side.

C3 - *rmm3* rDNA Has No DMS Reactive A at Nucleotide 1132 in Domain 2

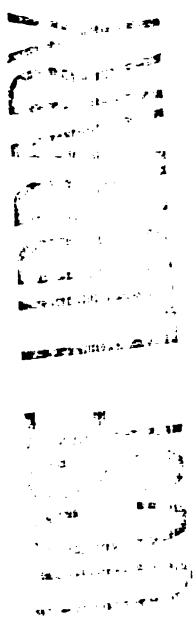


extended by a primer located from + 60 to + 40 towards the center of the molecule. The region seen on the gel extends from approximately - 10 to - 200. The brackets indicate the locations of the Type I repeats. The upstream borders of these repeats are labelled on the side above the corresponding downward facing arrow. rDNA nucleotide numbers are given on the side, adjacent to the A T sequencing ladder. Compare the - and + DMS lanes of the 8 minute timepoint. Footprinted residues are arrowed, and the nucleotide number of the site is indicated. There are several sites of G residue enhancements in the region of these two Type I repeats. These are located at - 20 from the start site of transcription at nucleotide 1867, within the Type Id repeat at nucleotide 1842, within the Type Ic repeat at nucleotides 1778 and 1779, and upstream of the Type Ic repeat at nucleotides 1761 and 1762. Sites at 1746 -1748 have been arrowed, although it is not clear whether these are truly sites of enhancement. At nucleotide 1792 there is a DMS reactive C residue, DMS methylates single-stranded Cs at the N-3 position, which is base paired in double-stranded DNA.

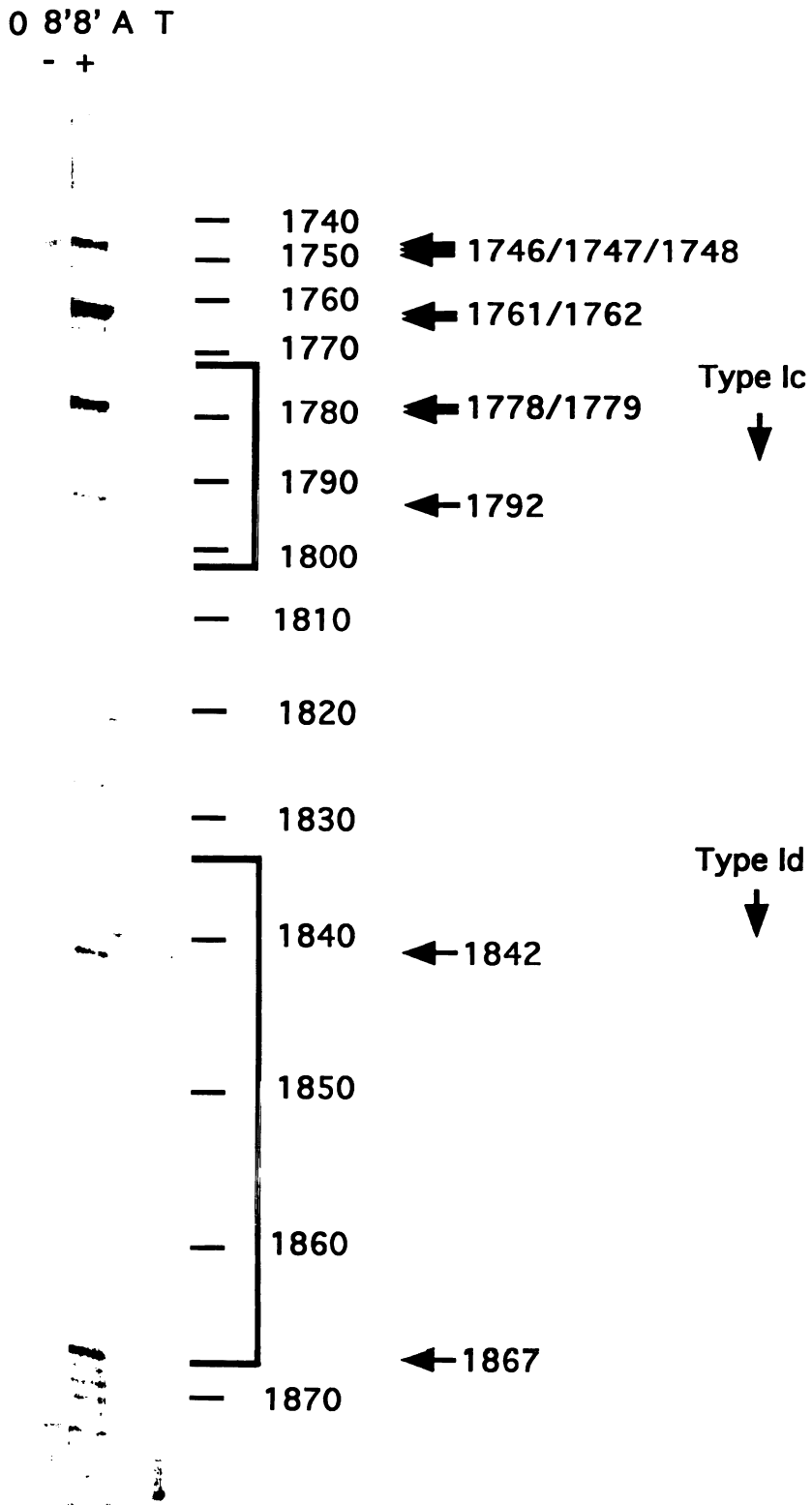
The *rmm3* base change is a deletion of an A in the central stretch of As in the Type Ic repeat. This A stretch falls between two footprinted sites, the reactive C at 1792 and the enhanced Gs at 1778 and 1779. The footprints at both sites are absent from *rmm3* strain cells (Figure 7b). In addition, the footprints at 1867 and 1842 are absent. However, the enhancement of nucleotide 1761 and perhaps 1762 is present. There do not appear to be footprinted sites

Figure 7a: C3 Wild-Type DMS Promoter Footprint

DMS treated DNA (see figure 6a) was primer extended with primer #12 at the rDNA promoter. This primer extends from within the trancription unit from + 60 towards the center of the molecule, detecting cleavages the coding strand of the promoter region. rDNA nucleotide numbers are indicated on the side, the positions of the Type Ic and d repeats are bracketed. Nucleotides that are enhanced in chromatin are indicated by the arrow. These are G residues, with the exception of nucleotide 1792 which is a C.



C3 Wild-Type
DMS Promoter Footprint



at 1746-1748 in *rmm3* strain cells. The DNA primer extended in Figure 7b is the same DNA that was primer extended in the Repeat 2 region in Figure 5b, and which had wild-type footprints at the Type III repeats and at nucleosome 5. Thus, in addition to the loss of the wild-type footprint at nucleotide 1132, *rmm3* strain cells have lost the wild-type footprint at the promoter region, with the exception of the enhancements upstream of the Type Ic repeat. Interestingly, in contrast to the loss of the wild-type footprint at the promoter region in the *rmm3* maintenance mutant, *rmm8* strain cells, which have a base change at -19 from the start site of transcription, have a nearly wild-type footprint at the promoter (Figure 8a and 8b). Thus, loss of the footprint at the promoter is not a general feature of promoter region chromosome maintenance mutants.

Discussion

We have described the identification of two new rDNA minichromosome maintenance mutants in *Tetrahymena*. *rmm7* has a base change in the Type Ib repeat of Domain 2. This finding confirms results from previous screens that the Type I repeat is important for chromosome maintenance. The previously reported mutants *rmm1* and *rmm4* were independent mutants identified in the same screen; each had a deletion of an A residue in the Type Ib repeat. That the *rmm7* mutation is an addition of an A in this repeat indicates that the effect on chromosome maintenance of these base changes results from the change in sequence at this site or to the change in spacing produced by a loss or a gain of an A residue. This

Figure 7b: C3-*rmm3* Lacks the Wild-Type DMS Promoter Footprint

DMS footprinting of the promoter in a C3-*rmm3* strain. The same DNA which was primer extended within Repeat 2 (see figure 4b) was primer extended with primer #12 at the promoter (see figure 7a). Note the absence of the enhancements seen in the wild-type with the exception of nucleotide 1761.

C3- *rmm3* Lacks the Wild-Type DMS Promoter Footprint

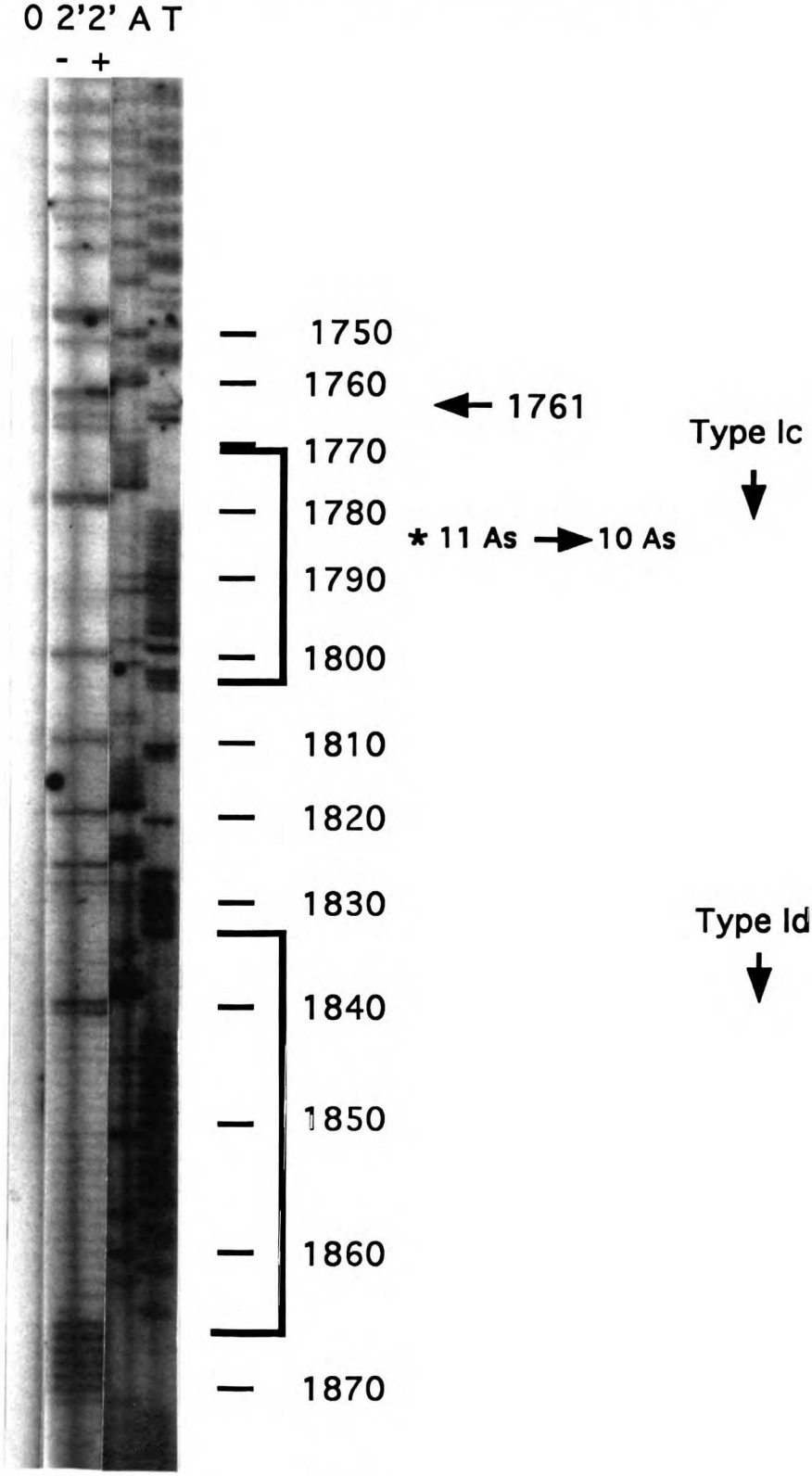


Figure 8a: C3 Wild-Type DNase I Promoter Footprint

DNase I footprinting at the promoter has been reported previously.

0 indicates naked DNA that has not been cleaved with DNase I

- indicates naked DNA control

+ indicates chromatin

.1, .2, 10 and 20 are the units/ml of DNase I in the reaction (total volume 200 microliters)

A and T are sequencing lanes. Plasmid was primer extended in the presence of dideoxy A or T.

Primer extension was done with primer #11 extension begins at - 200 and proceeds towards the transcription unit. rDNA nucleotide numbers are indicated along the side. The positions of the Type Ic and Id repeats are bracketed.

The start site of transcription is at nucleotide 1887.

C3- *rmm 8* Has a Nearly Wild-Type DNAse I Promoter Footprint

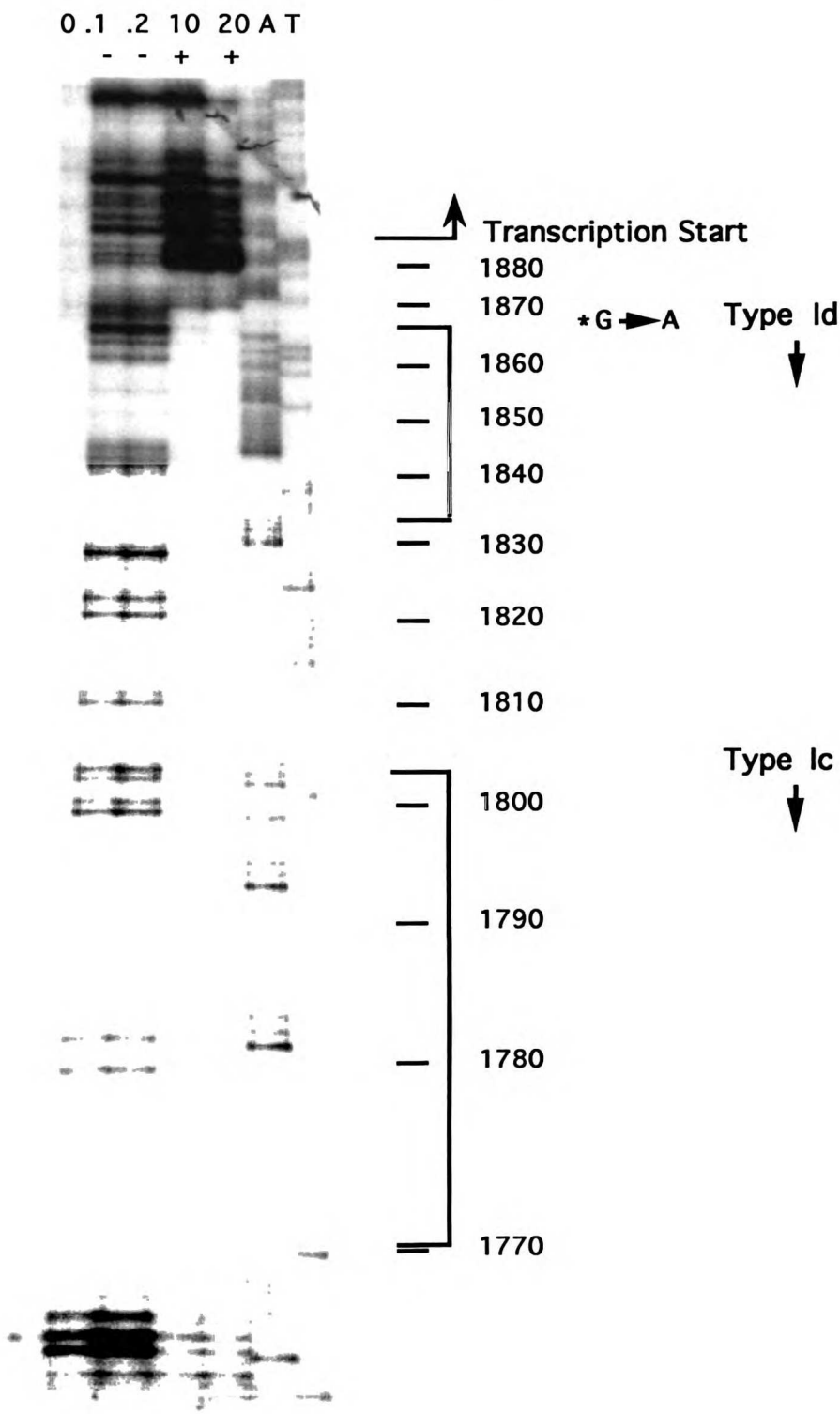
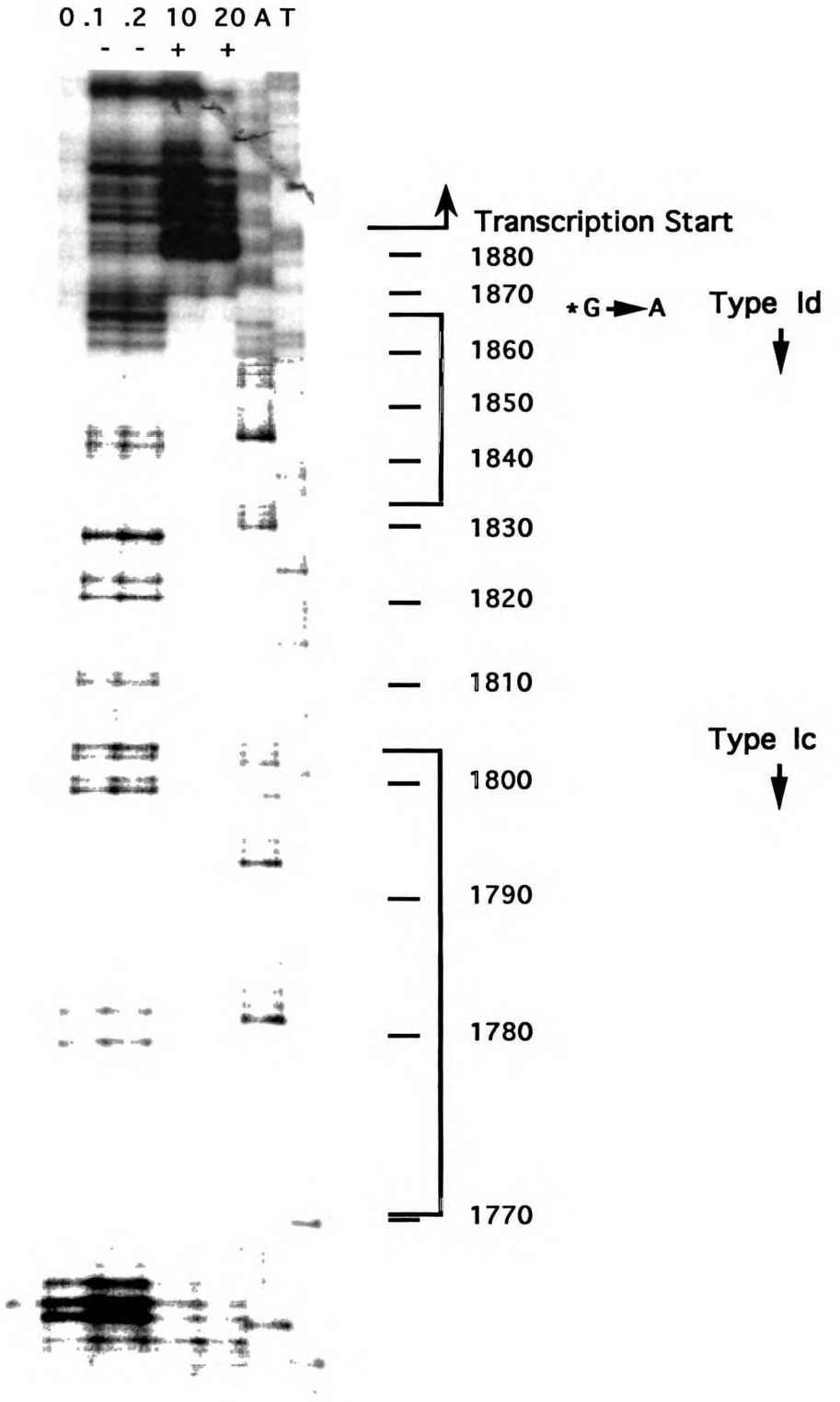


Figure 8b: C3-*rmm8* Has a Nearly Wild-Type DNase I Promoter Footprint

DNase I footprinting of SF112, a C3-*rmm8* maintenance mutant strain. The *rmm8* base change is at -19 from the promoter and is a G to A transition in a run of 6 G's, indicated by the asterisk.

C3- *rmm 8* Has a Nearly Wild-Type DNAse I Promoter Footprint



suggests that the base changes affect protein binding and not simply DNA structure; for example, a reduction in the number of A residues might have compromised a DNA unwinding element (DUE).

rmm8 is a G to A transition in the promoter region, at -19 from the start site of transcription. It is in a run of 6 Gs that begins with the terminal base of the promoter proximal Type Id repeat, and continues 3' toward the start site of transcription on the coding strand. It is in a larger sequence from -135 to +35 that has been conserved in Tetrahymenid ciliates (Challoner et al., 1985). Addition of a single G to this stretch of 6 Gs inactivates transcription from the Pol I promoter and results in the loss of the promoter footprint (Pan et al., 1995). Thus, this stretch of 6 G's is critical to promoter function. We have not tested promoter function in *rmm8* strain cells, but as shown above, *rmm8* strain cells do have a nearly wild-type DNase I footprint at the promoter. The cells have a slow-growth phenotype, but it has not been determined whether this is linked to the *rmm8* base change or not.

Three lines of evidence indicate that a wild-type promoter is important for wild-type chromosome maintenance. First, the identification of the *rmm8* maintenance mutant: reversion of the base change at the promoter to the wild-type sequence parallels the loss of the maintenance defect, suggesting that the G to A transition is responsible for the defect. Second, the observation that a base change in the promoter distal Type Ic repeat of *rmm3* strain cells causes a maintenance mutant phenotype (Orias and Larson, personal

communication; this work). However, the role of the Type Ic repeat in transcription has not been established, and, unlike the region mutated in *rmm8*, there is no direct evidence that it is important for transcription. Third, rDNA vectors transform cells more efficiently when they contain the promoter region, providing further evidence of the importance of this region for chromosome maintenance (Jacek Gaertig, personal communication).

In order to begin to understand the basis of the phenotypes observed in rDNA maintenance mutant chromosomes we chose to undertake a functional analysis of one of these, *rmm3*. This maintenance mutant has a base change in a Type I repeat, like *rmm1,4,* and *7*, and it is in the promoter region, like *rmm8*. Using neutral-neutral 2-D gels we demonstrated for the first time that an rmm mutant has a replication phenotype. We did not observe a transcription initiation phenotype. However, we observed the loss of footprints at the promoter and the origin region.

The observed *rmm3* replication phenotype is an enhanced accumulation of replicating intermediates within the 5' NTS that indicates that there is an increased replication fork pause. Such pauses have been seen in the rDNA of yeast, pea, Xenopus and human cells (Brewer et al., 1992; Hernandez et al., 1993; Hyrien et al., 1995; Little et al., 1993). The yeast replication fork barriers (RFB) occur at the 3' end of the upstream Pol I transcription unit and prevent replication from opposing transcription within the transcription unit. Thus, these rDNAs are replicated unidirectionally. In yeast the

RFB has been shown to act in a polar fashion, and to be due to protein binding at this site (Brewer and Fangman, 1988; Brewer et al., 1992). The barrier is independent of transcription. As mentioned above, human and *Xenopus,* have much larger 5' NTS regions and no single specific site of initiation. In these organisms there are multiple initiations at different sites on the same molecule, and numerous pauses are seen as a result of colliding forks. There has been speculation and experimentation about the effect of opposing RNA and DNA polymerases on the same molecule. There is evidence that transcription can have a profound effect on replication (Pan et al., 1995), and that it does not (French, 1992; Liu and Alberts, 1995; Liu et al., 1994; Liu et al., 1993). The effect may depend on the rate of transcription. Recently, Deshpande and Newlon, have demonstrated that transcription of a tRNA opposing movement of replication fork can produce a replication fork pause (Deshpande and Newlon, 1996).

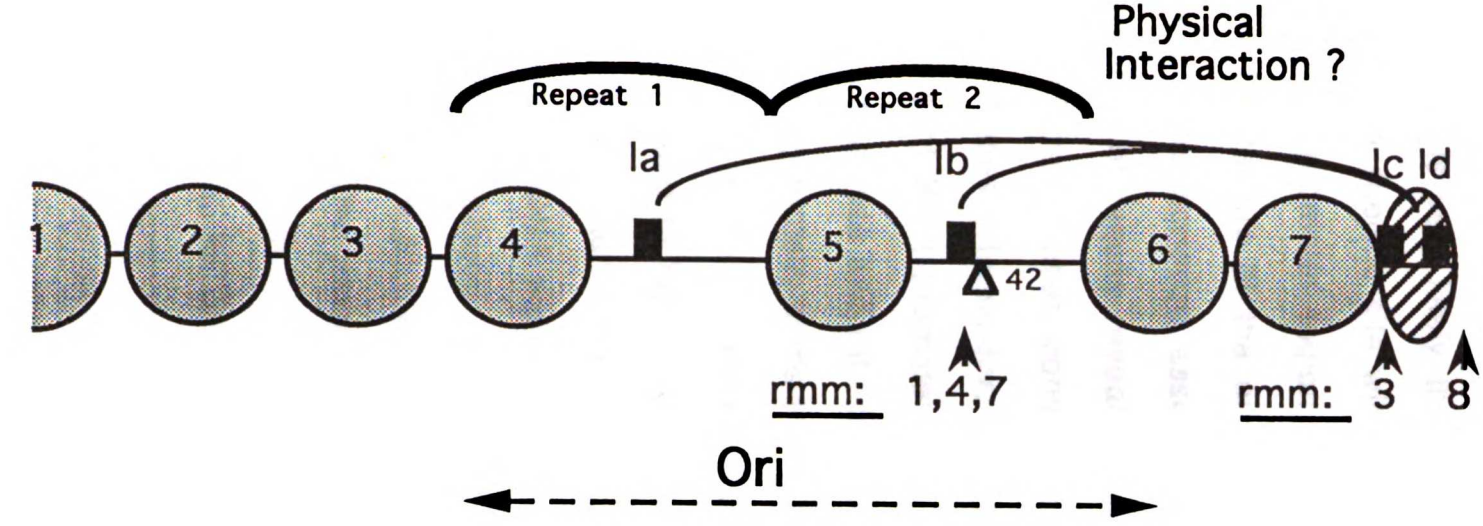
It is unlikely that the RFBs observed in the *Tetrahymena* rDNA are due to the act of transcription opposing the replication fork, since no stable transcripts have been observed in the 5' NTS. It is also unlikely that the RFBs are due to a fork that has crossed the center of the palindrome colliding with a fork that has initiated on that side, because of the proximity of one pause to the promoter, which is downstream of the replication origin (Kapler and Blackburn, unpublished results). It is most likely that the RFBs are due to a protein(s) binding the 5' NTS. The model that we favor is that a replication fork that has initiated on one side of the palindrome

collides with a structure on the other side of the palindrome that is established to initiate replication, and that involves interactions between the Domain 1, Domain 2, and promoter regions. These are the locations to which the RFBs have been mapped (Kapler and Blackburn, unpublished results), and these are the regions implicated by genetics and footprinting to be important in chromosome maintenance.

Our model is that a physical interaction between the Domain 1 and Domain 2 regions and the promoter is responsible for wild-type chromosome maintenance, perhaps in the manner described for the beta-globin locus in which there are stochastic interactions between the transcriptional enhancer locus control region and the promoter (Figure 9) (Wijgerde et al., 1995). This model is derived from our identification of chromosome maintenance mutants with base changes in the promoter region and in the Domain 2 region, and from our identification of the loss of wild-type footprints in the Domain 1, Domain 2, and promoter regions in a maintenance mutant (*rmm3*) with a base change in the promoter region. It is further supported by transformation results indicating that a wild-type promoter is necessary for optimal episomal maintenance (Jacek Gaertig, personal communication). In chapter two I showed that the majority of the 5' NTS in nuclei of log phase cells is occupied by seven highly positioned nucleosomes, and that the non-nucleosomal regions include the two Domains of Repeats 1 and 2 and the promoter region (Figure 8). Positioned nucleosomes have been shown to facilitate interactions between transcription factors to enhance

transcription (Schild et al., 1993). We propose that one function of the positioned nucleosomes of the 5' NTS may be to facilitate interactions between the intervening non-nucleosomal regions, and that nucleosomes may 6 and 7 wrap DNA in a way that allows the a protein complex at the promoter region to interact productively with a protein complex at the Domain 1/Domain 2 region. Further elaboration of this model will depend on identifying those protein complexes at the promoter and at the origin region and demonstrating directly that they interact.

Model of Interaction Important for Wild-Type rDNA Maintenance



Homologues of the origin recognition complex (ORC) proteins which bind the A element of the yeast ARS, have been identified in other organisms, to date none have been identified in *Tetrahymena*. (Carpenter et al., 1992). It is very likely that a complex similar to *S.cerevisiae* ORC binds at a specific site within the *Tetrahymena* rDNA origin region. In the promoter region it is virtually certain that one component of the factors at the origin is TATA-binding protein (TBP) and several transcription associated factors (TAFs). Several years ago a series of results revealed that TATA-binding protein is a component of transcription complexes at all three types of eukaryotic RNA polymerases (Hernandez, 1993). In the last several years the structure of Pol I promoters has been elucidated in higher organisms, *Xenopus*, mouse and rat. In addition to TBP plus TAFs, an HMG box containing protein binds these bipartite Pol I promoters (Heix and Grummt, 1995). However, *S.cerevisiae* seems to lack the HMG box protein and to have another complex of proteins present at the promoter (Keys et al., 1994). Thus, wild-type replication of the *Tetrahymena* rDNA appears to depend on an interaction between the Pol I promoter region and the origin, but it is not yet clear what proteins are at the Pol I promoter or what their specific role might be. The multiple roles transcription factors have been shown to play in viral and cellular origin function include altering DNA structure at an origin, recruiting core or auxiliary replication factors, helping to form a replication initiation complex, and chromatin remodeling (Heintz, 1992). This must be determined for each individual origin region.

Our model is based in part on our footprinting results for *rmm3*. We propose that this mutation in the promoter region weakens a structure involving interactions between a complex at the promoter and a complex at the origin region, resulting in the loss of a footprint at each location. However, this does not explain how the loss of such an interaction would result in the increased replication fork pause of *rmm3* strain cells. Determining this will entail identifying the nature of the pause sites and the structure of the complexes at the promoter and origin.

Current understanding of naturally occurring cellular origins of DNA replication is limited. The best studied origins are those of the yeast *Saccharomyces cerevisiae*, in which several functional origins have been well-defined, the protein complex which binds the consensus ARS sequence has been identified, and numerous genes involved in the control of DNA replication have been described (Diffley et al., 1995). Detailed analysis of the origin in *Tetrahymena* lags behind the work in yeast, but further study of this origin is likely to lead to insights relevant to all eukaryotes. Here we have used the unique features of this origin to uncover an apparent interplay between chromosomal regions that confers wild-type maintenance. This effect might not have been seen in an organism in which there was not some abrogation of cell cycle control, or in one in which chromosomes segregated faithfully. The unique features of the *Tetrahymena* rDNA have allowed this phenomenon to be uncovered.

We propose that an interaction between an upstream region and the promoter is important for wild-type chromosome maintenance. For the rDNA minichromosome in *Tetrahymena* the loss of this interaction can have a dramatic consequence, the loss of an entire allele of information, as is easily observed in a cross involving a chromosome with a drug resistance marker. However, it is the unusual features of the *Tetrahymena* macronuclear minichromosome which make this observation possible. The chromosomes is polyploid, there is no mechanism to insure faithful segregation of alleles, there is apparent abrogation of cell cycle regulation, and only one origin fires per molecule. In a more conventional nucleus with a chromosome that had multiple origins on a given molecule, and one in which faithful segregation of chromosomes has determined by the presence of a mitotic spindle, the interaction between the promoter complex and origin region would not have been observed. However, we believe that it is likely to be occurring in these nuclei as well.

Materials and Methods

Cell strains and culture:

Wild-type cells are C3 491-1a, a karyonide derived from the wild (Ed Orias). *rmm3* strain cells were obtained as heterokaryons from Ed Orias and were subjected to two rounds of mating in order to produce homozygous progeny. The strain used in these experiments is Round 2 #1. The slow growth phenotype of *rmm3* strain cells had been crossed out by Mio in the Orias lab. Confirmation that the homozygote contained the rmm3 base change was done by PCR of the

Type Ib repeat and sequencing of the PCR product. Cells were grown in 2% PPYS plus PSF at 30°C. Cells were starved by being washed twice in 1X Dryl's plus calcium and resuspended in this medium for between 16 and 24 hours. Cells were refed by the addition of 5% PPYS to 2%.

Generation of the rmm3 homozygote:

Backcrossing of the original mutagenized strain resulted in cells that lacked the slow growth phenotype and retained the *rmm3* base change. A heterokaryon was obtained from the laboratory of Ed Orias. This was crossed with an A* strain to produce homozygotes and these cells were allowed to mate with each other. This procedure should have produced cells homozygous for *rmm3* in the mic which expressed *rmm3* rDNA in the mac. The presence of the rmm3 base change was confirmed by PCR amplification and subsequent sequencing of the macronuclear rDNA and detection of the deletion of an A in the Type Ib repeat. These cells used in these studies were Rd2 #1.

Nuclear run-ons

Cell ghosts were prepared according to the procedure of Love et. al. (Love et al., 1988), and run-ons were performed according to the Red Book. Fifty mls of cells were grown to 2 x 10^5 cells/ml. Cells were chilled on ice and spun at 1500 rpm for five minutes in the HB-4 rotor. Cells were resuspended in extraction buffer plus 1% Triton X-100 and kept on ice for five minutes. Cells were then washed twice with extraction buffer at 4°C and resuspended in 250 l transcription

buffer 1. 200 l of cytoskeletal frameworks were incubated with 2 l 100 mM rNTPs (except U) and 32P-UTP (120 micro curies) for 2 minutes at 30°C in the presence of 100 l transcription buffer 2. Reactions were stopped with the addition of .6 ml HSB with 120 units of DNase I and were placed at 30°C for five minutes. .2 ml of SDS/Tris buffer with 1 mg/ml proteinase K was added and reactions were placed at 42°C for 30 minutes. Samples were extracted with phenol-chloroform, chloroform, and precipitated with ammonium acetate and isopropanol at - 20°C with the addition of 100 micrograms of tRNA. Precipitated sample was resuspended in 250 l DNAse Buffer with the addition of 30 units of DNase I per tube and incubated at 37°C for 30 minutes. 250 micrograms of proteinase K was added and the sample put at 42°C for thirty minutes. The sample was phenol-chloroform extracted, and 62.5 l 1 M NaOH was added to the aqueous and placed on ice for 10 minutes to hydrolyze the RNA to a uniform size. This was quenched with 125 l of 1 M Hepes, and precipitated with sodium acetate and ethanol. Samples were resuspended in water and an aliquot counted after washing on DE 81 paper. Equal numbers of counts from wild-type and mutant samples were added to 1.4 ml of hyb and hybridized at 42°C for 36 hours. Blots were washed at room temperature for 15 minutes is 2X SSC, and 60 minutes at 65°C in 2X SSC (2 x 30 minutes), then 30 minutes at 37°C in 2X SCC with 10 micrograms/ml RNAse, rinsed in 2X SSC at 37°C and autoradiographed at room temperature for 10 minutes. Filters were also placed on the phosphorImager for quantitation.

Two-Dimensional Gels:

Cells were grown in 200 ml cultures as described above. DNA was prepared by a modification of our usual procedure that includes reduced temperature of incubation and the substitution of proteinase K for Pronase. 200 mls of cells were spun down in a clinical centrifuge at 2,000 k rpm at room temperature for two minutes, resuspended in 10 - 15 mls of 10 mM Tris pH 7.5, re-spun and resuspended in 1 ml 10 mM Tris. 1 ml of NDS at 55°C was added and cells were placed at 37°C for 30 minutes. Then 40 l of 20mg/ml Proteinase K was added and incubation continued for 3 hours. Two mls of TE was added and samples were extracted with phenol-chloroform, phenol, and precipitated overnight at room temperature by the addition of 8 ml of ethanol. DNA was precipitated by spinning at 5,000 rpm in the HB-4 at 15°C, was washed, dried and resuspended in 1 ml TE with 10 micrograms/ml RNAse A and placed at 37°C for 45 minutes. DNA was reprecipitated with the addition of 1/10th volume 10 M NH4OAc and two volumes of ethanol. DNA was re-spun as above, washed dried and resuspended in 500 - 1000 l of TE. 100 - 200 micrograms of DNA was digested for three to six hours with a ten fold excess of enzyme to minimize loss of replicating intermediates. Samples were not processed with BND cellulose to enrich for replicating intermediates. Neutral-neutral gels were run essentially as reported by Brewer and Fangman (Brewer and Fangman, 1987). Hind III digests, containing a central 4.2 kb origin fragment were run in the first dimension on a .4 % agarose 1X TBE gel 25 cm x 15 cm with .1 micrograms of ethidium bromide/ml at 22 to 34 Volts for 24 to 17 hours at room

temperature. A 5 - 6 cm gel slice including the 1n and 2n sizes of the fragment of interest was excised from the gel under longwave UV light. The slice was rotated 90 and placed at the top of a 1% agarose 1X TBE gel, with .5 micrograms of ethidium bromide per ml, this solution was repoured around the gel slice, cooled, and run at 80 volts for 23 hours plus 130 for 4 hours in 1X TBE with .5 micrograms/ml EtBr at 4°C. XmnI I digests with a 1.6 kb origin containing fragment were run on .7 % agarose 1 X TBE gels in the first dimension at 22 volts for 24 hours, and on a 1.5 % agarose 1X TBE gel with .5 micrograms/ml EtBr in buffer with .5 micrograms/ml EtBr at 130 volts for 5.5 hours. Gels were photographed under short-wave UV light, depurinated for 10 - 15 minutes, denatured and transferred to Nytran by wicking in 10 X SSC for 12 - 24 hours. Kinased oligos in Church and Gilbert hybridization solution were hybridized at 45°C for 16 hours, washed at room temperature in C and G wash 100 mM and then at 45°C in C and G hyb 100 mM and autoradiographed for three to five days.

Genomic footprinting:

Was performed essentially as described in Pan et. al. (Pan et al., 1995). Nuclei were prepared as described in Palen and Cech (Palen and Cech, 1984). Cells were grown to a density of about 2 - 3 x 10 5 cells/ml spun down in a clinical centrifuge at 2,000 rpm for 2 minutes at room temperature, resuspended in 20 ml TMS with 1mM DTT, transferred to a cold sterile bottle, lysed in the cold by the addition of TMS containing 1% NP-40 to a final concentration of .16 %. Sucrose was added to 250 mM and allowed to dissolve in the cold

(about 20 minutes). Nuclei were pelleted from the lysed cells by spinning at 7,500 rpm in the HB-4 for thirty minutes at 4°C. Nuclei were resuspended in 800 l Buffer A with BME (except for DMS treatment) and divided into 100 l aliquots. For DNase I treatment cells were preincubated at 25°C for 2 minutes and DNase I was added for 1 minute. Reactions were stopped with the addition of 30 l .2 M EDTA and 200 l proteinase K solution, placed at 37°C overnight and phenol-chloroform extracted, chloroform extracted and precipitated and resuspended in TE. For KMnO4 reactions nuclei were pre incubated at 25°C for two minutes and then treated with 10 mM KMnO4 for various amounts of time. Reactions were stopped with the addition of BME stop and processed as above for DNase I. Treatment with DMS was also with 10 mM DMS for varying amounts of time, reactions were stopped with the addition of 9 volumes of .3 M BME in Buffer A. Nuclei were repelleted and lysed with proteinase K buffer and proteinase K as above. Naked DNA was prepared from nuclei prepared at the same time and was incubated with the reagent after removal of proteins from the DNA. Modified chromatin was compared to modified DNA by primer extension with Taq polymerase for the desired region of the rDNA. Samples were run on sequencing gels and exposed for two days to three weeks at - 80°C.

REFERENCES

Alfonso, P. J., Crippa, M. P., Hayes, J. J., and Bustin, M. (1994). The footprint of chromosomal proteins HMG-14 and HMG-17 on chromatin subunits [published erratum appears in J Mol Biol 1994 Jun 10;239(3):436]. *J Mol Biol* 236, 189.

Allis, C. D., Glover, C. V., Bowen, J. K., and Gorovsky, M. A. (1980). Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, Tetrahymena thermophila. *Cell* 20, 609.

Benard, M., Lagnel, C., and Pierron, G. (1995). Site-specific initiation of DNA replication within the non-transcribed spacer of Physarum rDNA. *Nucleic Acids Res* 23, 1447.

Bloom, K. S., and Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* 29, 305.

Bonven, B. J., Gocke, E., and Westergaard, O. (1985). A high affinity topoisomerase I binding sequence is clustered at DNAase I hypersensitive sites in Tetrahymena R-chromatin. *Cell* 41, 541.

Brewer, B. J., and Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell* 51, 463.

Brewer, B. J., and Fangman, W. L. (1988). A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* 55, 637.

Brewer, B. J., Lockshon, D., and Fangman, W. L. (1992). The arrest of replication forks in the rDNA of yeast occurs independently of transcription. *Cell* 71, 267.

Budarf, M. L., and Blackburn, E. H. (1986). Chromatin structure of the telomeric region and 3'-nontranscribed spacer of Tetrahymena ribosomal RNA genes. *J Biol Chem* 261, 363.

Bustin, M., Lehn, D. A., and Landsman, D. (1990). Structural features of th HMG chromosomal proteins and their genes. *Biologica and Biophysica Acta* 1049, 231.

Buttinelli, M., Di, M. E., and Negri, R. (1993). Multiple nucleosome positioning with unique rotational setting for the Saccharomyces cerevisiae 5S rRNA gene in vitro and in vivo. *Proc Natl Acad Sci U S A* 90, 9315.

Carpenter, P. B., Mueller, P. R., and Dunphy, W. G. (1992). Role for a Xenopus Orc2-related protein in controlling DNA replication. *Mol Cell Biol* 12, 4305.

Cartwright, I. L., and Kelly, S. E. (1991). Probing the Nature of Chromosomal DNA-Protein Contacts by *In Vivo* Footprinting. *Biotechniques* 11, 188.

Cech, T. R., and Brehm, S. L. (1981). Replication of the extrachromosomal ribosomal RNA genes of Tetrahymena thermophilia., 3531.

Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E., and Campbell, J. L. (1984). Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae. *Mol Cell Biol* 4, 2455.

Challoner, P. B., Amin, A. A., Pearlman, R. E., and Blackburn, E. H. (1985). Conserved arrangements of repeated DNA sequences in nontranscribed spacers of ciliate ribosomal RNA genes: evidence for molecular coevolution. *Nucleic Acids Res* 13, 2661.

Cheng, L., and Kelly, T. J. (1989). Transcriptional activator nuclear factor I stimulates the replication of SV40 minichromosomes in vivo and in vitro. *Cell* 59, 541.

DePamphilis, M. L. (1993). Eukaryotic DNA replication: anatomy of an origin. *Annu Rev Biochem* 62, 29.

Deshpande, A. M., and Newlon, C. S. (1996). DNA Replication Fork Pause Sites Dependent on Transcription. *Science* 272, 1030.

Diffley, J. F., and Cocker, J. H. (1992). Protein-DNA interactions at a yeast replication origin. *Nature* 357, 169.

Diffley, J. F., Cocker, J. H., Dowell, S. J., Harwood, J., and Rowley, A. (1995). Stepwise assembly of initiation

complexes at budding yeast replication origins during the cell cycle. *Journal of Cell Science* Supplement 19, 67.

Diffley, J. F., Cocker, J. H., Dowell, S. J., and Rowley, A. (1994). Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell* 78, 303.

Diffley, J. F. X., and Stillman, B. (1990). The initiation of chromosomal DNA replication in eukaryotes. *TIG* 6, 427.

Elliot, A. M. (1973). Biology of *Tetrahymena* (Stroudsbug, Pennsylvania: Dowden, Hutchinson and Ross, Inc.).

Engberg, J., Din, N., Saiga, H., and Higashinakagawa, T. (1984). Nucleotide sequence of the 5'-terminal coding region for pre-rRNA and mature 17S rRNA in Tetrahymena thermophila rDNA. *Nucleic Acids Res* 12, 959.

Engberg, J., Mowat, D., and Pearlman, R. E. (1972). Preferential Replication of the Ribosomal RNA Genes During a Nutritional Shift-Up in *Tetrahymena pyriformis*. *Biochemica et Biophysica Acta* 272, 312.

Engberg, J., and Nielsen, H. (1990). Complete sequence of the extrachromosomal rDNA molecule from the ciliate *Tetrahymena thermophila* strain B1868VII. *Nucleic Acids Research* 18, 6915.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. *Nature* 355, 219.

Frappier, L., and O'Donnell, M. (1992). EBNA1 distorts oriP, the Epstein-Barr virus latent replication origin. *J Virol* 66, 1786.

French, S. (1992). Consequences of replication fork movement through transcription units in vivo. *Am J Clin Oncol* 15, 41.

Gaertig, J., and Gorovsky, M. A. (1995). DNA-mediated transformation in Tetrahymena. *Methods Cell Biol* 47, 559.

Gall, J. G. (1986). The Molecular Biology of Ciliated Protozoa (Orlando, Florida: Academic Press, Inc.).

Gilson, E., Roberge, M., Giraldo, R., Rhodes, D., and Gasser, S. M. (1993). Distortion of the DNA double helix by RAP1 at silencers and multiple telomeric binding sites. *J Mol Biol* 231, 293.

Gorovsky, M. A. (1986). Ciliate Chromtin and Histones. *The Molecular Biology*, 227.

Hay, R. T., and DePamphilis, M. L. (1982). Initiation of SV40 DNA replication in vivo: location and structure of 5' ends of DNA synthesized in the ori region. *Cell* 28, 767.

Hayatsu, H., and Ukita, T. (1967). The Selective Degredation of Pyrimidines in Nucleic Acids by Permanganate Oxidation. *Biochemical and Biophysical Research Communications* 29, 556.

Hayes, J., J., Clark, D. J., and Wolffe, A. P. (1991). Histone contributions to the structure of DNA in the nucleosome. *Proc. Natl. Acad. Sdi. USA* 88, 6829.

Hayes, J. J. (1995). Chemical probes of DNA structure in chromatin. *Chemistry and Biology* 2, 127.

He, Z., Brinton, B. T., Greenblatt, J., Hassell, J. A., and Ingles, C. J. (1993). The transactivator proteins VP16 and GAL4 bind replication factor A. *Cell* 73, 1223.

Heintz, N. H. (1992). Transcription factors and the control of DNA replication. *Curr Opin Cell Biol* 4, 459.

Heix, J., and Grummt, I. (1995). Species specificity of transcription by RNA polymerase I. *Current Opinion in Genetics & Development* 5, 652.

Hernandez, N. (1993). TBP, a universal eukaryotic transcription factor? *Genes Dev* 7, 1291.

Hernandez, P., Martin, P. L., Martinez, R. M., and Schvartzman, J. B. (1993). Conserved features in the mode of replication of eukaryotic ribosomal RNA genes. *Embo J* 12, 1475.

Hogan, M. E., Rooney, T. F., and Austin, R. H. (1987). Evidence for kinks in DNA folding in the nucleosome. *Nature* 328, 554.

Huberman, J. A., Spotila, L. D., Nawotka, K. A., el, A. S., and Davis, L. R. (1987). The in vivo replication origin of the yeast 2 microns plasmid. *Cell* 51, 473.

Hyrien, O., Maric, C., and Mechali, M. (1995). Transition in specification of embryonic metazoan DNA replication origins. *Science* 270, 994.

Jakobovits, E. B., Bratosin, S., and Aloni, Y. (1980). A nucleosome-free region in SV40 minichromosomes. *Nature* 285, 263.

Kapler, G. M., and Blackburn, E. H. (1994). A weak germ-line excision mutation blocks developmentally controlled amplification of the rDNA minichromosome of Tetrahymena thermophila. *Genes Dev* 8, 84.

Kapler, G. M., Orias, E., and Blackburn, E. H. (1994). Tetrahymena thermophila mutants defective in the developmentally programmed maturation and maintenance of the rDNA minichromosome. *Genetics* 137, 455.

Keys, D. A., Vu, L., Steffan, J. S., Dodd, J. A., Yamamoto, R. T., Nogi, Y., and Nomura, M. (1994). RRN6 and RRN7 encode subunits of a multiprotein complex essential for the initiation of rDNA transcription by RNA polymerase I in Saccharomyces cerevisiae. *Genes Dev* 8, 2349.

Kister, K. P., Kaffenberger, W., and Eckert, W. A. (1988). In vitro synthesis and processing of pre-rRNA in isolated macronuclei from Tetrahymena. *Eur J Cell Biol* 46, 233.

Kister, K. P., Muller, B., and Eckert, W. A. (1983). Complex endonucleolytic cleavage pattern during early events in the processing of pre-rRNA in the lower eukaryote, Tetrahymena thermophila. *Nucleic Acids Res* 11, 3487.

Larson, D. D., Blackburn, E. H., Yaeger, P. C., and Orias, E. (1986). Control of rDNA replication in Tetrahymena involves a cis-acting upstream repeat of a promoter element. *Cell* 47, 229.

Larson, D. D., Umthun, A. R., and Shaiu, W.-L. (1991). Copy Number Control in the *Tetrahymena* Macronuclear Genome. *J. Protozool.* 38, 258.

Li, R., and Botchan, M. R. (1993). The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate in vitro BPV-1 DNA replication. *Cell* 73, 1207.

Little, R. D., Platt, T. H., and Schildkraut, C. L. (1993). Initiation and termination of DNA replication in human rRNA genes. *Mol Cell Biol* 13, 6600.

Liu, B., and Alberts, B. M. (1995). Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science* 267, 1131.

Liu, B., Wong, M. L., and Alberts, B. (1994). A transcribing RNA polymerase molecule survives DNA replication without aborting its growing RNA chain. *Proc Natl Acad Sci U S A* 91, 10660.

Liu, B., Wong, M. L., Tinker, R. L., Geiduschek, E. P., and Alberts, B. M. (1993). The DNA replication fork can pass RNA polymerase without displacing the nascent transcript. *Nature* 366, 33.

Love, H. J., Allen, N. A., Zhao, Q. A., and Bannon, G. A. (1988). mRNA stability plays a major role in regulating the temperature-specific expression of a Tetrahymena thermophila surface protein. *Mol Cell Biol* 8, 427.

Lutter, L. C. (1978). Kinetic analysis of deoxyribonuclease I cleavages in the nucleosome core: evidence for a DNA superhelix. *J Mol Biol* 124, 391.

Marahrens, Y., and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255, 817.

Maxam, A. M., and Gilbert, W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci U S A* 74, 560.

McGhee, J. D., and Felsenfeld, G. (1979). Reaction of nucleosome DNA with dimethyl sulfate. *Proc Natl Acad Sci U S A* 76, 2133.

McPherson, C. E., Shim, E. Y., Friedman, D. S., and Zaret, K. S. (1993). An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. *Cell* 75, 387.

Meersseman, G., Pennings, S., and Bradbury, M. (1991). Chromatosome Positioning on Assembled Long Chromatin

Linker Histones Affect Nucleosome Placement on 5 S rDNA. *J. Mol. Biol.* 220, 89.

Miller, C. A., and Kowalski, D. (1993). cis-acting components in the replication origin from ribosomal DNA of Saccharomyces cerevisiae. *Mol Cell Biol* 13, 5360.

Mukhopadhyay, G., Carr, K. M., Kaguni, J. M., and Chattoraj, D. K. (1993). Open-complex formation by the host initiator, DnaA, at the origin of P1 plasmid replication. *Embo J* 12, 4547.

Nawotka, K. A., and Huberman, J. A. (1988). Two-dimensional gel electrophoretic method for mapping DNA replicons. *Mol Cell Biol* 8, 1408.

Orias, E., and Bradshaw, A. D. (1992). Stochastic developmental variation in the ratio of allelic rDNAs among newly differentiated, heterozygous macronuclei of Tetrahymena thermophila. *Dev Genet* 13, 87.

Orlando, V., and Paro, R. (1995). Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr Opin Genet Dev* 5, 174.

Palen, T. E., and Cech, T. R. (1984). Chromatin structure at the replication origins and transcription-initiation regions of the ribosomal RNA genes of Tetrahymena. *Cell* 36, 933.

Pan, W. J., and Blackburn, E. H. (1995). Tandem repeats of the 5' non-transcribed spacer of Tetrahymena rDNA function as high copy number autonomous replicons in the

macronucleus but do not prevent rRNA gene dosage regulation. *Nucleic Acids Res* 23, 1561.

Pan, W. J., Gallagher, R. C., and Blackburn, E. H. (1995). Replication of an rRNA gene origin plasmid in the Tetrahymena thermophila macronucleus is prevented by transcription through the origin from an RNA polymerase I promoter. *Mol Cell Biol* 15, 3372.

Paull, T. T., Haykinson, M. J., and Johnson, R. C. (1993). The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes Dev* 7, 1521.

Pennings, S., Meersseman, G., and Bradbury, E. M. (1991). Mobility of Positioned Nucleosomes on 5 S rDNA. *J. Mol. Biol.* 220, 101.

Prescott, D. M. (1994). The DNA of ciliated protozoa. *Microbiol Rev* 58, 233.

Prioleau, M. N., Huet, J., Sentenac, A., and Mechali, M. (1994). Competition between chromatin and transcription complex assembly regulates gene expression during early development. *Cell* 77, 439.

Pruss, D., Bushman, F. D., and Wolffe, A. P. (1994). Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc Natl Acad Sci U S A* 91, 5913.

Reichel, R., and Benecke, B. J. (1984). Localization of small nuclear polymerase I RNA sequences at the 5' end of the human rDNA transcription unit. *Embo J* 3, 473.

Sasse-Dwight, S., and Gralla, J. D. (1990). Role of Eukaryotic-Type Functional Domains Found in the Prokaryotic Enhancer Receptor Factor σ^{54}. *Cell* 62, 945.

Schild, C., Claret, F. X., Wahli, W., and Wolffe, A. P. (1993). A nucleosome-dependent static loop potentiates estrogen-regulated transcription from the Xenopus vitellogenin B1 promoter in vitro. *Embo J* 12, 423.

Schulman, I. G., Wang, T., Wu, M., Bowen, J., Cook, R. G., Gorovsky, M. A., and Allis, C. D. (1991). Macronuclei and Micronuclei in Tetrahymena thermophila Contain High-Mobility-Group-Like Chroomosomal Proteins Containing a Highly Conserved Eleven-Amino-Acid Putative DNA-Binding Sequence. *Molecular and Cellular Biology* 11, 166.

Shen, X., Yu, L., Weir, J. W., and Gorovsky, M. A. (1995). Linker histones are not essential and affect chromatin condensation in vivo. *Cell* 82, 47.

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

Simpson, R. T. (1992). Nucleosome Positioning: Occurrence, Mechanisms, and Functional Consequences. *Progress in Nucleic Acid Research and Molecular Biology* 40, 143.

Simpson, R. T., and Stafford, D. W. (1983). Structural features of a phased nucleosome core particle. *Proc Natl Acad Sci U S A* 80, 51.

Simpson, R. T., and Whitlock, J. P. (1976). Mapping DNAase I-susceptible sites in nucleosomes labeled at the 5' ends. *Cell* 9, 347.

Stargell, L. A., Bowen, J., Dadd, C. A., Dedon, P. C., Davis, M., Cook, R. G., Allis, C. D., and Gorovsky, M. A. (1993). Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila. *Genes Dev* 7, 2641.

Stargell, L. A., and Gorovsky, M. A. (1994). TATA-binding protein and nuclear differentiation in Tetrahymena thermophila. *Mol Cell Biol* 14, 723.

Thoma, F., Bergman, L. W., and Simpson, R. T. (1984). Nuclease digestion of circular TRP1ARS1 chromatin reveals positioned nucleosomes separated by nuclease-sensitive regions. *J Mol Biol* 177, 715.

Truett, M. A., and Gall, J. G. (1977). The replication of ribosomal DNA in the macronucleus of Tetrahymena. *Chromosoma* 64, 295.

van de Wetering, M., and Clevers, H. (1992). Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *Embo J* 11, 3039.

Walker, S. S., Francesconi, S. C., and Eisenberg, S. (1990). A DNA replication enhancer in *Saccharomyces cerveisiae*. *Proc. Natl. Acad. Sci.* 87, 4665.

Wang, T., and Allis, C. D. (1993). An abundant high-mobility-group-like protein is targeted to micronuclei in a cell cycle-dependent and developmentally regulated fashion in Tetrahymena thermophila. *Mol Cell Biol* 13, 163.

Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. *Nature* 377, 209.

Wolffe, A. P. (1991). Activating chromatin. *Current Biology* 1, 366.

Wolffe, A. P. (1995). Centromeric chromatin. Histone deviants. *Curr Biol* 5, 452.

Wolffe, A. P. (1994). Transcriptional activation. Switched-on chromatin. *Curr Biol* 4, 525.

Wolffe, A. P., and Dimitrov, S. (1993). Histone-modulated gene activity: developmental implications. *Crit Rev Eukaryot Gene Expr* 3, 167.

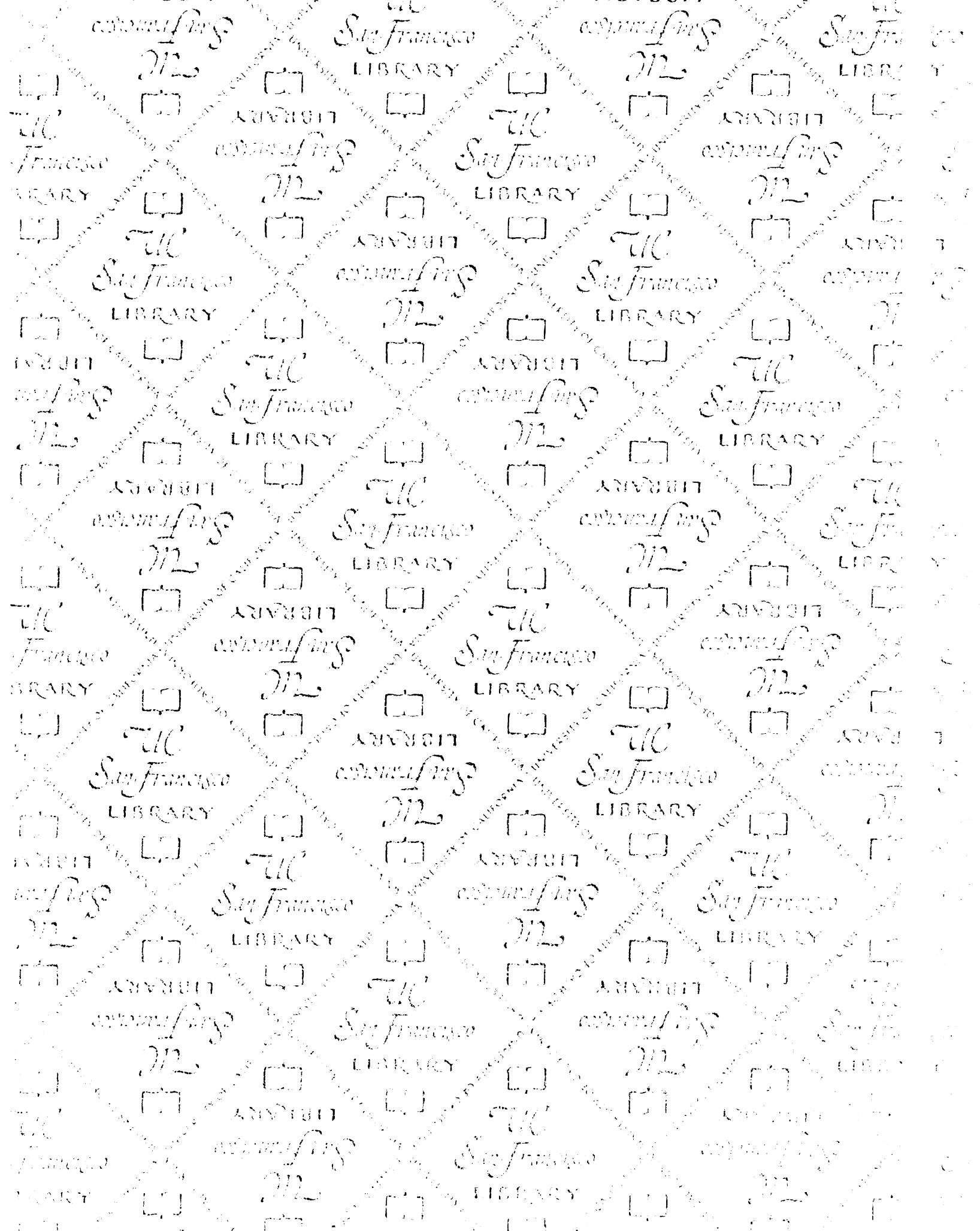
Yaeger, P. C., Orias, E., Shaiu, W. L., Larson, D. D., and Blackburn, E. H. (1989). The replication advantage of a free linear rRNA gene is restored by somatic recombination in Tetrahymena thermophila. *Mol Cell Biol* 9, 452.

Yang, L., and Botchan, M. (1990). Replication of Bovine Papillomavirus Type 1 DNA Initiates within an E2-Responsive enhancer Element. *Journal of Virology* 64, 5903.

Yu, S.-M., Horowitz, S., and Gorovsky, M. A. (1987). A novel approach for studying gene expression in the cell cycle reveals coordinate and independent regulation of members of the H4 multigene family in cycling and in nongrowing Tetrahymena. *Genes and Development* 1, 683.

Zhang, L., and Gralla, J. D. (1989). In situ nucleoprotein structure at the SV40 major late promoter: melted and wrapped DNA flank the start site. *Genes Dev* 3, 1814.

Zhao, K., Kas, E., Gonzalez, E., and Laemmli, U. K. (1993). SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *Embo J* 12, 3237.


The San Francisco LIBRARY