1993

Characterization of cis-acting replication control elements in ribosomal RNA genes of Tetrahymena thermophila

Wen-Ling Shaiu

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Genetics Commons, Molecular Biology Commons, and the Zoology Commons

Recommended Citation

https://lib.dr.iastate.edu/rtd/10858

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Characterization of cis-acting replication control elements in ribosomal RNA genes of *Tetrahymena thermophila*

Shaiu, Wen-Ling, Ph.D.
Iowa State University, 1993
Characterization of cis-acting replication control elements in ribosomal RNA genes of *Tetrahymena thermophila*

by

Wen-Ling Shaiu

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of DOCTOR OF PHILOSOPHY

Department: Zoology and Genetics
Major: Zoology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1993

Copyright © Wen-Ling Shaiu, 1993. All rights reserved.
DEDICATION

This dissertation is dedicated to my father and mother, who have influenced me greatly and taught me the philosophy of life.
# TABLE OF CONTENTS

## GENERAL INTRODUCTION
- Rationale 1
- Tetrahymena Biology 4
- Studies of the replication origin in rDNA 5
- Genetic analysis of Tetrahymena 6
- Explanation of dissertation format 8

## PAPER I: THE TYPE I REPEAT IN RIBOSOMAL rRNA GENES OF *TETRAHYMENA* IS A CIS-ACTING REPLICATION CONTROL ELEMENT

## ABSTRACT

## INTRODUCTION

## MATERIALS AND METHODS
- Cell culture and Tetrahymena strains 19
- DNA isolation 19
- Southern blot analysis 19
- Cloning and DNA sequence analysis 19

## RESULTS
- Identification of cis-acting elements affecting rDNA replication 21
- Rare somatic recombination events lead to a prevalence of rDNAs with C3-rmm replication capabilities in C3-rmm4/B-Pmr mass culture 22
- The rDNA from long-term C3-rmm4/B-Pmr cultures lacks the 1-bp deletion of C3-rmm4 rDNA and the 42-bp deletion of the B-Pmr rDNA 23

## DISCUSSION
REFERENCES

PAPER II: ANALYSIS OF CONSERVED FEATURES IN THE REPLICATION ORIGIN REGION OF TETRAHYMENA rDNA

ABSTRACT

INTRODUCTION

MATERIALS AND METHODS

Computer analysis of Tetrahymena rDNA sequences
Culture and rDNA isolation
Bent DNA analysis
Isolation of scaffold-associated DNA
Two-dimensional gel electrophoresis

RESULTS

Common modular sequence elements clustered within the NTSs
An intrinsically bent DNA segment is located in the origin-containing fragment
A scaffold-associated region is located near the palindromic center
Analysis of replicative intermediates by two-dimensional gel electrophoresis

DISCUSSION

REFERENCES

PAPER III: ATOMIC FORCE MICROSCOPY OF ORIENTED LINEAR DNA MOLECULES LABELED WITH 5NM GOLD SPHERES

ABSTRACT

INTRODUCTION

MATERIALS AND METHODS

DNA labeling
Sample deposition and AFM imaging
GENERAL INTRODUCTION

Rationale

DNA replication is highly regulated to ensure that the genome is duplicated accurately in both prokaryotes and eukaryotes. A critical difference between prokaryotic and eukaryotic modes of DNA replication is that prokaryotic chromosomes comprise a single replicon while eukaryotic genomes are much larger and consist of multiple replicons. How the initiation of replication at multiple sites on large eukaryotic chromosomes is regulated and coordinated is not known. In fact, very little is known about the DNA sequences that constitute origins of replication in eukaryotes. A major goal of this work is to define specific sequence elements required for activity of a known eukaryotic origin of replication, the replication origin in the rRNA genes (rDNA) of *Tetrahymena thermophila*.

DNA sequences and protein factors involved in initiation of DNA replication in prokaryotes have been well studied (review in Marians 1992). In prokaryotes, initiation of DNA replication occurs at specific sequences called origins of replication. The process of replication begins with the recognition of origin-specific sequences by initiator proteins and the formation of DNA-protein complexes that induces structural changes in the duplex DNA at the origin (Bramhill and Kornberg 1988). Other sequence domains near or within each origin region have intrinsically low helical stabilities and readily undergo localized unwinding in response to initiator protein binding. This is followed by association of a series of replication proteins at the partially single-stranded region. DNA synthesis begins at structures called replication forks, where unwinding of the double helix is coupled with chain elongation both temporally and spatially. Thus in bacteria, the mode of DNA replication is known as the replication fork mechanism.
Although the replication fork mechanism is often invoked to explain DNA replication in eukaryotic chromosomes, only two features of this mechanism have ever been demonstrated: the presence of short nascent fragments (Okazaki-type fragments), and the presence of replication micro-bubbles (Benbow et al. 1986; Gaudette and Benbow 1986; Umek et al. 1989). Moreover, examination of replicating DNA from *Xenopus* and several other eukaryotes failed to reveal replication "bubbles" at the frequency expected (Benbow et al. 1986). Instead, large regions of single-stranded DNA have been observed in samples of DNA derived from S-phase cells (Baldari et al. 1978; Bjursell et al. 1979; Gaudette and Benbow 1986; Henson 1978; Wortzman and Baker 1980). Therefore, a strand separation mechanism has been suggested as an alternative model for chromosomal DNA replication in eukaryotes (Benbow 1985; Benbow et al. 1986; Gaudette and Benbow 1986). In this model, DNA synthesis may be uncoupled from DNA unwinding, and extensive regions of single-stranded DNA are observed frequently. DNA synthesis can be initiated at many sites throughout both single-stranded templates, resulting in a mixture of Okazaki-type fragments and long nascent DNA chains appearing on both DNA strands (Benbow 1985; Gaudette and Benbow 1986; Méchali and Kearsey 1984).

The DNA sequence elements that function as origins of replication in eukaryotes are not well characterized, except in yeast and mammalian viral origins, e.g., in SV40 (DePamphilis 1987; Hay and DePamphilis 1982), adenovirus (Challberg and Kelly 1979), Epstein-Barr virus (Gahn and Schildkraut 1989; Reisman et al. 1985), Bovine papilloma Virus (Schvartzman et al. 1990; Waldeck et al. 1984), and polyomavirus (Hendrickson et al. 1987; Muller et al. 1983). The only eukaryotic definitive chromosomal origins of replication are those associated with autonomously replicating sequences (ARSs) in the yeast, *Saccharomyces cerevisiae*. ARSs were originally isolated based on their ability to direct the autonomous replication of plasmids in yeast (Stinchcomb et al. 1979; Struhl et al. 1979). All ARSs share several structural features. Each ARS has a 100-200 bp long AT-rich sequence that contains an 11 bp ARS core consensus sequence (ACS) (review in Campbell and Newlon 1991; Fangman and Brewer 1991;
Newlon 1988). The ACS is indispensable for the function of the DNA origin (Marahrens and Stillman 1992; Van Houten and Newlon 1990), and is the binding site for a multiprotein complex, the origin recognition complex (ORC) (Bell and Stillman 1992; Diffley and Cocker 1992). Deletions in the flanking sequences of ACS reveal several additional functional and structural domains, including binding sites for transcriptional factors, DNA unwinding elements, and bent DNA sequences (Anderson 1986; Buchman et al. 1988; Celniker et al. 1984; Kearsey 1984; Marahrens and Stillman 1992; Shore and Nasmyth 1987). Although ARS function has been primarily studied on plasmids, recent work has shown that some but not all ARSs can function as origins in their normal chromosomal context (Brewer and Fangman 1991; Deshpande and Newlon 1992; Ferguson and Fangman 1992; Huberman et al. 1988; Linskens and Huberman 1988; Walker et al. 1991; Zhu et al. 1992). Activity of the chromosomal origin has been shown to be dependent on the ARS core consensus sequence (Deshpande and Newlon 1992; Walker et al. 1991). It is not yet known what additional sequences are needed.

Genomic DNA fragments from other organisms, including Tetrahymena (Amin and Pearlman 1986; Luehrsen et al. 1988), Drosophila (Gragerov et al. 1988; Mills et al. 1986; Stinchcomb et al. 1980), human (Montiel et al. 1984), plants (Eckdahl and Anderson 1989), and others (Stinchcomb et al. 1980) can also function as ARSs in yeast. However, there is no evidence that these DNA sequences that behave as ARSs in yeast also function as origins of replication in their native context. Nevertheless, by examining DNA sequences, ARS-like sequences also have been found near several eukaryotic origins of replication (Benbow et al. 1992; Hernández et al. 1993; Liang et al. 1993).

Many experimental approaches for identification of origins of DNA replication have been developed. A comprehensive summary of these methods describing the rationale, advantages, and limitations has been reviewed (Vassilev and DePamphilis 1992). These approaches are divided into three groups: (1) methods based on analysis of newly synthesized DNA strands
including fiber autoradiography, analysis of early-replicating fragments, replication origin trap, nascent strand extrusion, nascent DNA strand length, and replication-fork polarity; (2) methods based on analysis of DNA structures, including electron microscopy (EM), and two-dimensional gel electrophoresis (2-D gel mapping); (3) methods based on functionality, including assays of ARS activity in yeast or in mammalian cells. Putative origins of replication have been reported in many organisms using some of the mapping techniques mentioned above. Table 1 shows a representative list of putative chromosomal replication origins from fourteen eukaryotic organisms.

We are investigating the sequence requirements for eukaryotic replication origin function by focusing on the amplified rDNA of *Tetrahymena thermophila*. This ciliated protozoan offers the advantages of a known origin of replication (which has been roughly mapped) in a molecule that is experimentally tractable, both genetically and biochemically. The remainder of this chapter provides background concerning the structure and organization of the *Tetrahymena* rDNA and summarizes previous genetic and molecular studies of rDNA replication.

*Tetrahymena* rRNA genes

*Tetrahymena thermophila* is a ciliated protozoan. Each *Tetrahymena* cell contains two structurally and functionally differentiated nuclei: a germline micronucleus and a somatic macronucleus. The micronucleus is diploid and transcriptionally quiescent; it functions to provide gamete nuclei for the sexual process. The macronucleus is polyploid and transcriptionally active; it determines the phenotype of the cell. Both nuclei are derived from a single zygotic nucleus formed during mating, and, thus, are initially identical. However, the developing macronucleus undergoes a wide spectrum of chromosomal alterations, including chromosomal fragmentation, diminution and polyploidization of 90% of micronuclear genomic sequences to 45C, with specific amplification of the rDNA to 10,000 copies (Blackburn and Karrer 1986).
Of the processes associated with macronuclear development, maturation of the rDNA is the most extensively studied (Yao 1986). A great deal is known about the molecular processes of rDNA in the two nuclei and some of the steps involved in the pathway leading to rDNA amplification (Figure 1). In the micronucleus, each haploid genome contains a single, chromosomally-integrated copy of the rDNA, located on the left arm of chromosome II. By contrast, the mature vegetative macronucleus contains 10,000 extrachromosomal rDNA molecules. The extrachromosomal rDNA is a linear 21 kb long molecule, which is made up of two inverted repeats of the micronuclear rDNA sequence. In developing macronuclei, a linear 11 kb molecule containing a single copy of the rDNA has been identified (Pan and Blackburn 1981). The sites at which chromosome breakage and sequence elimination at each end of the micronuclear copy occur to generate free linear molecules, have been determined (Yao 1986; Yao 1989; Yao and Yao 1989; Yao et al. 1990), but the processes leading to palindrome formation and the role of the 11 kb molecule are not understood. In each half of the palindromic molecule, about 2 kb of DNA sequence makes up the 5' nontranscribed spacer (5'NTS). The 3' nontranscribed spacer (3'NTS), also about 2 kb in size, terminates in variable numbers of telomeric repeats, \((G_4T_2)_n\), averaging 400 bp in length.

Studies of replication origin in rDNA

Cech and Brehm mapped an origin of replication approximately 600 bp from the center of the rDNA of *T. thermophila* by electron microscopy (Cech and Brehm 1981). The location of the origin was based on the presumption that the center of a replication bubble (or eye) should be the origin, and that the two replication forks progress at the same rate. The rDNA molecules vary in length, due to different numbers of telomeric repeats at their termini. Therefore, taking the heterogeneity of telomere length into account, the replication origin of rDNA was estimated to map at 650 bp ± 300 bp from the palindromic center (Palen and Cech 1984).
Restriction fragments from both 5' and 3' NTS regions of the rDNA exhibit ARS activity in yeast (Amin and Pearlman 1985; Amin and Pearlman 1986). Interestingly, one of these ARS-containing 5'NTS fragments roughly coincides with the origin of replication as mapped by electron microscopy. In addition, circular plasmids containing two tandem arrays of the 5'NTS and one copy of the rRNA coding sequence can be maintained and replicate autonomously after microinjection into *Tetrahymena* macronuclei (Yu and Blackburn 1989; Yu *et al.* 1988). However, this plasmid often recombines with the 5'NTS of the endogenous rDNA and generates molecules that exhibit a replication advantage. In other words, the products of recombination bearing multiple copies of 5'NTS exhibit a replicative advantage over the non-recombinant endogenous molecules (Yu and Blackburn 1990). A recombinant molecule with more copies of the 5'NTS exhibits a replication advantage over molecules with fewer copies of the 5'NTS. The results from these studies suggest strongly that the 5'NTS region contains the origin of replication.

**Genetic analysis of *Tetrahymena* rDNA**

A genetic selection scheme was developed to isolate mutants defective in amplification and/or replicative maintenance of the rDNA in the macronucleus and identified a cis-acting replication control element (Larson *et al.* 1986; Yaeger *et al.* 1989). Two previous findings were exploited in these studies. First, a dominant mutation, *Pmr* that confers resistance to the protein synthesis inhibitor paromomycin, is located in the 17S rRNA gene of inbred strain B and provides a selectable marker within the rDNA (Spangler and Blackburn 1985). Second, a naturally occurring variant allele of the rDNA found in inbred strain C3 is distinguishable from that found in inbred strain B based on its replication potential and restriction site polymorphisms (Pan *et al.* 1982). In C3 and B crosses, greater than 99% of the heterozygous progeny contain only C3 rDNA in their mature macronuclei (Larson *et al.* 1986). In other
words, C3 rDNA predominates over B rDNA in the macronucleus, although the two alleles are inherited in the micronucleus in a Mendelian fashion.

Several *rmm* (*rDNA maturation and maintenance*) mutants bearing mutations in the C3 rDNA, were isolated in the laboratory of Dr. E. Orias at University of California, Santa Barbara. These mutations were detected based on their ability to abolish the dominance of C3 rDNA over B rDNA in heterozygotes. Molecular and development studies of these mutants revealed a replication hierarchy in which C3 > B > C3-*rmm* in replicative ability (Larson et al. 1986).

Figure 2 illustrates the structure of the macronuclear palindromic rDNA. There are three repeated sequence elements (Types I, II, and III repeats) that are conserved in the 5'NTS among several ciliate species (Challoner et al. 1985; Niles et al. 1981). Some of these elements are clustered in two regions designated Domain 1 and Domain 2. These two domains are nuclease hypersensitive and have been shown to be free of nucleosomes (Bonven et al. 1985; Palen and Cech 1984). The Type III repeats have been demonstrated to be topoisomerase I binding and cleavage sites (Bonven et al. 1985; Christiansen et al. 1987). The role of the Type II repeats is under investigation by another group. The Type I repeat proximal to the transcriptional start site has been shown to be an essential part of the promoter for rRNA transcription (Miyahara et al. 1993).

DNA sequencing of the 5'NTS of C3, B, and one of the *rmm* mutants (C3-*rmm1*) revealed a 42-base-pair (bp) deletion in B rDNA, and a 1-bp deletion within *rmm1* rDNA compared with C3 rDNA. Both deletions affected the Type I repeat element in Domain 2 of the 5'NTS (Figure 2C). The 1-bp deletion occurs in a tract of 11 A nucleotide residues (Larson et al. 1986), and has a more pronounced effect than the 42-bp deletion on the replication property of rDNA, probably because it changes the critical spacing of conserved sequences within the Type I repeat (Larson et al. 1986). These studies provided the first genetic evidence for specific *cis*-acting replication control sequences in *Tetrahymena*. 
The goal of the work described in this dissertation is to define the Tetrahymena rDNA origin of replication. Studies in several eukaryotes have suggested that chromosomal origins of replication are considerably larger and more complex than their prokaryotic counterparts. The experiments described here are directed at testing this hypothesis.

Explanation of dissertation format

This dissertation presents studies of the cis-acting control sequences for rDNA replication in *Tetrahymena*, including genetic studies and physical mapping experiments. In addition, a powerful new technique, atomic force microscopy, was employed in an attempt to develop a high resolution physical mapping method for the rDNA replication origin and other eukaryotic replication origins.

There are four papers in this dissertation. Paper I consists of two sets of results: (1) My work on the determination of the sequence alteration in a revertant rDNA allele derived from a somatic recombination event. The work has been published in Molecular and Cellular Biology (Yaeger, P. C., Orais, E., Shaiu, W.-L., Larson, D. D., and Blackburn, E. H. Mol. Cell Biol. 9: 452-460. 1989). (2) My molecular analysis of the C3-rmm3 mutant, which suggested a regulatory link between replication and transcription of the rRNA genes. This work is the basis of a manuscript in preparation.

Paper II presents an analysis of common modular sequence elements in the rDNA of *Tetrahymena* and mapping of the *in vivo* rDNA replication origin by two-dimensional gel electrophoresis.

Table 1
Chromosomal origins of DNA replication in eukaryotes

<table>
<thead>
<tr>
<th>Organisms</th>
<th>Genes</th>
<th>Estimated size</th>
<th>Mapping Methods*</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Physarum</em></td>
<td>rDNA</td>
<td>4 kb</td>
<td>a,b</td>
<td>(Daniel and Johnson 1989; Vogt and Braun 1977)</td>
</tr>
<tr>
<td>polycephalum</td>
<td>profilin</td>
<td>1.5 kb</td>
<td>b</td>
<td>(Bénard and Pierron 1992)</td>
</tr>
<tr>
<td><em>Sea urchin</em></td>
<td>rDNA</td>
<td>70-100 kb</td>
<td>a</td>
<td>(Botchan and Dayton 1982),</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>rDNA</td>
<td>571 bp</td>
<td>a,b</td>
<td>(Linskens and Huberman 1988; Saffer and Miller 1986)</td>
</tr>
<tr>
<td><em>S. pombe</em></td>
<td>ARS1 (TRPI)</td>
<td>193 bp</td>
<td>b,h</td>
<td>(Ferguson et al. 1991; Marahrens and Stillman 1992)</td>
</tr>
<tr>
<td></td>
<td>ARS121 (ch.10)</td>
<td>489 bp</td>
<td>b,h</td>
<td>(Walker et al. 1991)</td>
</tr>
<tr>
<td></td>
<td>ura4 (ch.3)</td>
<td>3-5 kb</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td><em>T. thermophila</em></td>
<td>rDNA</td>
<td>2 kb</td>
<td>a</td>
<td>(Cech and Brehm 1981)</td>
</tr>
<tr>
<td><em>T. pyriformis</em></td>
<td>rDNA</td>
<td>1 kb</td>
<td>a</td>
<td>(Truett and Gall 1977)</td>
</tr>
<tr>
<td><em>Sciara</em></td>
<td>DNA puff</td>
<td>6 kb</td>
<td>b</td>
<td>(Liang et al. 1993)</td>
</tr>
<tr>
<td><em>coprophila</em></td>
<td>II/9A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Drosophila</em></td>
<td>chorion genes</td>
<td>0.9-12 kb</td>
<td>a,b</td>
<td>(Delidakis and Kafatos 1989; Heck and Spradling 1990; McKnight and Miller 1977)</td>
</tr>
<tr>
<td><em>melanogaster</em></td>
<td>histone gene</td>
<td>10-15 kb</td>
<td>b</td>
<td>(Shinomiya and Sawako 1991)</td>
</tr>
<tr>
<td><em>Xenopus</em></td>
<td>rDNA</td>
<td>70-100 kb</td>
<td>a</td>
<td>(Bozzi et al. 1981)</td>
</tr>
<tr>
<td><em>laevis</em></td>
<td>a-globin genes</td>
<td>1.7 kb</td>
<td>e</td>
<td>(Razin et al. 1986)</td>
</tr>
</tbody>
</table>

* (a) genomic Southern hybridization, (b) in vivo labeling, (c) isolation and characterization of oriC DNA, (d) in vitro initiation of DNA replication, (e) nick translation labeling, (f) aDNA puffing, (g) histone H1 binding, (h) plasmid integration. 
<table>
<thead>
<tr>
<th>CHO cells</th>
<th>DHFR gene</th>
<th>0.45-28 kb</th>
<th>b,d,f</th>
<th>(Burhans et al. 1986; Burhans et al. 1990)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rhodopsin gene</td>
<td>5 kb</td>
<td>d</td>
<td>(Gale et al. 1992)</td>
</tr>
<tr>
<td>Mouse</td>
<td>Adenosine Deaminase gene</td>
<td>11 kb</td>
<td>f,g</td>
<td>(Carroll et al. 1993)</td>
</tr>
<tr>
<td>Human</td>
<td>c-myc gene</td>
<td>2.4 kb</td>
<td>c (in Hela cells)</td>
<td>(Ariga et al. 1989; McWhinney and Leffak 1990)</td>
</tr>
<tr>
<td></td>
<td>L30E of ch.19</td>
<td>2 kb</td>
<td>d</td>
<td>(Vassilev and Johnson 1990)</td>
</tr>
<tr>
<td></td>
<td>rDNA</td>
<td>0.19 kb</td>
<td>b,f</td>
<td>(Van't Hof et al. 1987)</td>
</tr>
<tr>
<td>Pea (Pisum sativum)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* a: EM; b: 2-D gel mapping; c: ARS activity; d: nascent strand size analysis; e: nascent strand extrusion; f: early-replicating fragments; g: replication-fork polarity; h: mutational analysis of functional domains

Note: Fourteen ARS elements have been identified and mapped in chromosome III of *S. cerevisiae* (Newlon et al. 1991). Six of these ARS elements have been shown to serve as chromosomal origins by 2-D gel mapping. Due to the long list of all references, they are not listed in the table. However, they are reviewed in the work of Greenfeder and Newlon (Greenfeder and Newlon 1992) as well as in Dershowitz and Newlon (Dershowitz and Newlon 1993).
Figure 1. Diagram of the molecular processes in the pathway of rDNA amplification in *Tetrahymena thermophila*. Forms of the rDNA in the micronucleus, the developing macronucleus, and the mature vegetative macronucleus are shown. The solid blocks, thick lines, and thin dashed lines designate the transcribed regions, nontranscribed spacers, and chromosomal sequences surrounding the micronuclear rRNA gene, respectively. Vertical bars represent telomeric elements (G₄T₂). Maturation events include excision of the micronuclear rRNA gene from flanking sequences (Yao 1986); acquisition of mature telomeric structures; self-replication of 11 kb rDNA molecules; conversion of the single gene into a palindrome; amplification of extrachromosomal palindromic rDNA molecules; and replication and maintenance of rDNA molecules at 10,000 copies in the vegetative macronucleus.
Figure 2. Structure of *Tetrahymena* palindromic rDNA and comparison of 5'NTS sequences of C3, B, and C3-rmm1. (A) Map of the 21 kb palindromic molecule with the axis of symmetry indicated by a dashed line and telomeric (TTGGGG)n sequences represented by hatched lines (■). (B) The big arrow heads designate nuclease hypersensitive regions redrawn from Palen and Cech (1984). The origin of replication (ori) (Cech and Brehm, 1981) and Domain 1 (D1) and Domain 2 (D2) are shown. (C) Organization of the repeated elements in the 5'NTS of three rDNA alleles. The positions of 1-bp and 42-bp deletions are indicated. Type I repeat, ■. Type II repeat, □. Type III, □.
THE TYPE I REPEAT IN RIBOSOMAL rRNA GENES OF *TETRAHYMENA* IS
A *CIS*-ACTING REPLICATION CONTROL ELEMENT
THE TYPE I REPEAT IN RIBOSOMAL rRNA GENES OF TETRAHYMENA IS A CIS-ACTING REPLICATION CONTROL ELEMENT

Wen-Ling Shaiu

Drena D. Larson

Department of Zoology and Genetics,
Iowa State University, Ames, IA 50011

This paper consists of two sets of data. The first set of data concerning somatic recombination has been published (Yaeger, P. C., Orais, E., Shaiu, W.-L., Larson, D. D., and Blackburn, E. H. Mol. Cell Biol. 9: 452-460. 1989). My contribution included all the cloning and sequencing analysis. In the second set of data, which describes another rDNA allele, C3-rrn3, I sequenced the mutant rDNA and identified the location of mutation.
ABSTRACT

The autonomously replicating rRNA genes (rDNA) in the somatic nucleus of *Tetrahymena thermophila* are maintained at a copy number of approximately 10,000 per nucleus. Independently-derived mutants, C3-rmm3 and C3-rmm4, in which the replication properties of the rDNA are altered, were isolated and characterized. In both mutants, the rDNA replication defect was associated with a single base pair (bp) deletion within a copy of the Type I repeat sequence, a conserved repeated element located in the 5' non-transcribed spacer of the rDNA (5'NTS) (Larson et al. 1986). The C3-rmm3 mutation alters a copy of the Type I repeat associated with the transcriptional promoter region, whereas the C3-rmm4 mutation is located about 600 bp upstream from the transcriptional start site. The wild-type C3-rDNA allele has a replication advantage over the rDNA allele found in inbred strain B, which has been shown to contain a 42 bp deletion. This deletion is separated by 25 bp from the 1 bp deletion of C3-rmm4. During prolonged maintenance of vegetative cultures of C3-rmm4/B heterozygotes, somatic recombination produced rDNA molecules lacking both the *rmm4*-associated 1-bp deletion and the 42-bp deletion. In somatic nuclei in which this rare recombinational event had occurred, all 10,000 copies of nonrecombinant rDNA were eventually replaced by the recombinant rDNA. These results prove that the deletions identified in these alleles are the genetic determinants of the observed replication phenotypes. Furthermore, they provide genetic evidence for a *cis*-acting control element that may play a dual role in DNA replication and transcription.
INTRODUCTION

The rDNA of the ciliated protozoan, *Tetrahymena thermophila*, provides an excellent system for studying eukaryotic DNA replication. Each *Tetrahymena* cell contains two structurally and functionally differentiated nuclei: a germ line micronucleus and a vegetative macronucleus. The macronuclear rDNA is relatively short (21 kb), autonomously replicating molecule that is maintained at high copy number (about \(10^4\)). In the germline micronucleus, the rDNA is present as a single copy per haploid genome, making it amenable to genetic analysis. Both the macronucleus and micronucleus are mitotic descendants of a single zygotic nucleus. The micronuclei in the mating parental cells are the precursors of the macronuclei in their progeny.

Previous work has shown that rDNA from inbred strain C3 (C3 rDNA) has a replicative advantage over rDNA from inbred strain B (B rDNA), causing C3 rDNA to completely replace B rDNA in C3/B heterozygotes during vegetative growth (Larson et al. 1986). Both types of rDNA were amplified early in macronuclear development in these heterozygotes. Specifically, a mixture of these rDNAs was present in populations of heterozygous cells up to at least 40 generations after conjugation. These results suggested that preferential replication of C3 rDNA in dividing cells was the major factor accounting for the finding that macronuclei of vegetatively maintained heterozygotes contained exclusively C3 rDNA (Larson et al. 1986).

Although previous molecular studies have identified the processes of rDNA excision, palindrome formation, telomere acquisition, amplification, and maintenance of high copy number, the mechanisms by which these processes occur are not understood. Moreover, the regulatory interactions that govern these processes have not been identified. Therefore, a genetic analysis is critical for a full understanding of this system.

A genetic scheme was developed to isolate *cis*-acting mutations, henceforth referred to as *rmm* mutations, that affect rDNA maturation or maintenance in the macronucleus (Larson et al.
18

1986). Using this scheme, several such \textit{rmn} mutants have been isolated. Among them, the C3-
\textit{rmn1} has been studied in detail. Briefly, C3-\textit{rmn1} was derived from mutagenesis of strain C3,
and C3-\textit{rmn1} rDNA allele has lost its ability to outreplicate the B rDNA allele in C3-\textit{rmn1}/B
heterozygotes. In this paper, two other mutants (C3-\textit{rmn3} and C3-\textit{rmn4}) exhibiting a similar
phenotype were analyzed. The results support and extend the findings from C3-\textit{rmn1}, namely
that a specific \textit{cis}-acting control element, the Type I repeat, is involved in the regulation of
rDNA replication in \textit{Tetrahymena}.

Despite the fact that B rDNA outreplicates \textit{rmn} rDNA in both C3-\textit{rmn1}/B-\textit{Pmr} and C3-
\textit{rmn4}/B-\textit{Pmr} heterozygotes (\textit{Pmr} designates paromomycin, the antibiotic selective marker
located on the 17S rRNA gene of B rDNA), long term vegetative cultures of these
heterozygotes sometimes appear to contain C3 rather than B rDNA in the macronucleus, as
determined by restriction fragment length polymorphism analysis (Yaeger \textit{et al.} 1989). The
revertant rDNAs were cloned and sequenced to characterize this unusual finding. In the
analysis of C3-\textit{rmn4}/B-\textit{Pmr} heterozygotes, rare somatic events were shown to yield
recombinant rDNAs capable of replacing the majority species of rDNA in a population of
heterozygotes. Characterization of these recombinant rDNAs provides unequivocal proof that
the 1-bp deletion (associated with C3-\textit{rmn1} and C3-\textit{rmn4}) and the 42-bp deletion (within B
rDNA) are the determinants for the replication disadvantage of these rDNA alleles. The
significance of these genetic changes will be discussed.
MATERIALS AND METHODS

Cell culture and *Tetrahymena* strains. All cultures were grown and maintained in 2% PPYS medium (Larson *et al.* 1986), supplemented with drugs when specified. Mating and other routine genetic procedures were as previously described (Orias 1986).

DNA isolation. rDNA molecules from mass cultures of C3-*rmn4*/B-*pmr* and C3-*rmn3*/B-*pmr* heterozygotes were isolated by the whole cell preparation procedure of Din and Engberg (Din and Engberg 1979) with modifications described by Spangler and Blackburn (Spangler and Blackburn 1985).

Southern blot analysis. Detection of C3 rDNA versus B rDNA was based on their restriction fragment length polymorphism. BamHI and SphI were used to distinguish C3 from B rDNA. An extra BamHI site exists in the 3'NTS of C3 rDNA and an extra SphI site in the 5'NTS of B rDNA. To confirm the complete restriction enzyme digestion of the rDNA, pUC18 DNA was included in the BamHI and SphI digestions. After electrophoresis in an agarose gel, DNA was transferred to Nytran filters (S&S) in 0.4M NaOH (Reed and Mann 1985) to minimize intramolecular renaturation of the palindromic central rDNA fragments. Nick-translated 32P-labeled DNA probes were hybridized to filters by standard procedures (Sambrook *et al.* 1989). The probes used were pTtR1 and p6J, in which segments of the rDNA 5' or 3'NTSs have been inserted (Challoner *et al.* 1985).

Cloning and DNA sequence analysis. The source of pure C3-*rmn4* and C3-*rmn3* rDNA was from crosses of strain SB841 (C3-*rmn4*) and SB837 (C3-*rmn3*) with strain CU374, which lacks the left arm of chromosome 2 in the micronucleus and, therefore, has no micronuclear form of the rDNA. These crosses produced hemizygous cells containing only C3-*rmn4* or C3-*rmn3* rDNA in their macronuclei. Macronuclear rDNA was purified from these cells and prepared for cloning into pUC119 by digestion either with XbaI alone or a combination of XbaI and TaqI (Yaeger *et al.* 1989). Single-stranded DNA was produced from
these clones by superinfection of the bacterial host strain MV1193 with helper phage M13KO7 (Vieira and Messing 1987). Dideoxy sequencing was performed according to the manufacturer's description (Pharmacia; USB). Sequence analysis of C3-rmm3 was performed using both universal and reverse primers (New England Biolabs) to examine sequence change from both XbaI and TaqI sites. To confirm the dideoxy sequencing results, C3-rmm3 clones were reexamined by Maxam and Gilbert chemical sequencing.

For DNA sequence analysis of rDNA molecules generated in C3-rmm1/B-pmr and C3-rmm4/B-Pmr mass cultures, rDNA was isolated from the cultures after long-term vegetative growth and the 820-bp XbaI-TaqI fragment of the 5'NTS was cloned into the plasmid pUC119. The DNA sequence of the region containing the 1-bp deletion in C3-rmm1 rDNA and the 42-bp deletion in B rDNA was verified by analysis of four independently cloned isolates from each recombinant rDNA preparation.
RESULTS

Identification of a cis-acting element affecting rDNA replication. C3-rmm4 was one of several mutants that affected maturation and maintenance of rDNA in macronuclei. The characterization procedures for C3-rmm4 were identical to those described for C3-rmm1 (Larson et al. 1986). C3-rmm4 reversed the replication advantage of wild-type C3 rDNA, causing the loss of C3 rDNA in macronuclei initially heterozygous for C3 and B rDNA. This was shown by mating a strain homozygous in the germ line for C3 rDNA and C3-rmm4 (C3-rmm4) with a strain homozygous for B rDNA and Pmr (B-Pmr). Pmr is a dominant mutation within the rDNA that confers resistance to the antibiotic paromomycin (Spangler and Blackburn 1985). Individual cell lines from the population of newly formed C3-rmm4/B-pmr heterozygotes were maintained and challenged with paromomycin. All of the cell lines remained Pmr-resistant throughout long-term vegetative growth, demonstrating a persistence of the B-Pmr rDNA. Molecular analysis of the cell lines, using restriction fragment length polymorphisms between C3 and B rDNA in the 5'NTS and 3'NTS, confirmed that the C3 rDNA had been replaced by B rDNA (Yaeger et al. 1989).

All of the characteristics reported above for C3-rmm4 are the same as those reported for C3-rmm1. Furthermore, the sequence analysis showed that C3-rmm4 is associated with the same 1-bp deletion within the 5'NTS as that previously described for C3-rmm1 (Figure 1).

The isolation scheme for C3-rmm3 was the same as that used for C3-rmm1 and C3-rmm4. C3-rmm3 shows a very similar replication deficiency phenotype to C3-rmm1 and C3-rmm4. The sequence analysis of C3-rmm3 revealed that this mutant also has a single base pair deletion within a Type I repeat element, however, in this case, the mutation is located in Type I repeat upstream of the site of transcriptional initiation (Figure I; Yaeger et al. 1989).

This genetic lesion contributes to the failure to maintain C3 type rDNA in C3-rmm3/B-Pmr heterozygotes, again strongly suggesting that the Type I repeat is involved in the
mechanism that regulates the replication of rDNA molecules. Moreover, a copy of Type I repeat element has been demonstrated to be an essential element of the rRNA gene promoter in *in vitro* transcription experiments using extracts of *Tetrahymena thermophila* cells (Miyahara *et al.* 1993). This finding has a significant implication that RNA transcription and DNA replication may share common regulatory elements in *Tetrahymena* cells.

**Rare somatic recombination events lead to a prevalence of rDNAs with C3-rmm replication capabilities in C3-rmm4/B-Pmr mass cultures.** The determination of which type of rDNA exists in the vegetative heterozygous C3-rmm4/B-Pmr and C3-rmm1/B-Pmr cultures was based on restriction fragment length polymorphisms (RFLPs) in the 5' and 3'NTSs of C3 and B rDNA, as well as the antibiotic selection marker Pmr on B rDNA (Figure 2). DNA isolated from individual cell lines at different time points was tested for the presence of the 3'NTS BamH I site in C3 (referred to as 3'BamH-C3) and the 5'NTS Sph I site in B rDNA by Southern blot analysis. At the earliest time point tested, both C3 and B rDNA were present, indicating that C3-C3-rmm4 and B rDNA had been amplified in the developing macronucleus. The amount of C3 type rDNA gradually decreased over 80 generations. However, after 90 fissions, a reversal of this trend occurred, such that the 3'BamH-C3 was detected in the population during the next 200 fissions. This pattern and timing of the disappearance of Pmr-resistance phenotype were almost identical in the results from three experiments. The results from the testing of single-cell isolates from these cultures for Pmr-resistance were consistent with the data from the 3'BamH-C3 analysis. The resurgence of the 3'BamH-C3 in the cultures suggested that somatic recombination between rDNA molecules was involved. Therefore, further molecular analysis of the 5'NTS Sph I site of B rDNA (referred to as 5'Sph-B) was performed. Analysis of restriction digestions of the revertant rDNA exhibited 5'Sph-B. These data indicate the existence of a 5'Sph-B and 3'BamH-C3 recombinant rDNA (Yaeger *et al.* 1989).
The rDNA from long-term C3-rmm4/B-Pmr cultures lacks the 1-bp deletion of C3-rmm4 rDNA and the 42-bp deletion of the B-Pmr rDNA. The prevalence of recombinant rDNA molecules in C3-rmm4/B-Pmr cultures raised the possibility that recombination events had "reverted" the mutation responsible for reduced replication capability of C3-rmm4 mutant. To test this possibility, rDNA from a C3-rmm4/B-Pmr mass culture was purified and cloned into pUC119.

Sequences of four independently cloned isolates, containing the XbaI-TaqI fragment distal to the palindromic center, were examined in the region containing Domain 2 (see Figure 1), where the rmm4-associated 1-bp deletion is located. Domain 2 is a near perfect duplication of Domain 1 (which roughly corresponds to the EM-mapped replication origin). The sequence results showed that Domain 2 of the recombinant molecules is in fact indistinguishable from that of wild-type C3 rDNA (Figure 3A). Thus, the crossover event generating these rDNAs must have occurred within the 25-bp region that separates the 1-bp deletion of C3-rmm4 rDNA and the 42-bp deletion of B-Pmr rDNA (Figure 4A).

In the C3-rmm1/B-Pmr mass cultures, no recombinant rDNA molecules were detected by restriction analysis. However, the similarity in the phenotype, i.e., the percentage of Pmr sensitive cells increased with time (Yaeger et al. 1989), suggested that rDNA molecules with wild type C3 rDNA replication capabilities could have been generated in the C3-rmm1/B-Pmr cultures by a reversion or recombination event that was not detected by RFLP analysis. Therefore, rDNAs from this long term mass culture were purified and sequenced. The Domain 2 region of four cloned isolates containing the XbaI-TaqI fragment was examined. In contrast to the results obtained for the C3-rmm4/B-Pmr culture, rDNA from C3-rmm1/B-Pmr culture still contained the 1-bp deletion characteristic of the C3-rmm1 mutant. Examination of about 200 bp of the Domain 2 region revealed no changes from the previously determined mutant C3-rmm1 rDNA sequence (Figure 3B). Thus, the resurgence of C3 type rDNA in the
heterozygous culture cannot be explained simply on the basis of restoration of Domain 2 sequences as in the C3-rmm4/B-Pmr cultures.
DISCUSSION

Two independent mutants, C3-<i>rmm1</i> and C3-<i>rmm4</i>, displayed the same phenotype and DNA sequence alteration, and were both shown to affect rDNA replication (Larson et al. 1986; Yaeger et al. 1989). The finding of 1-bp deletion in the Type I repeat of two independent mutants strongly argues that these mutations are indeed responsible for the phenotypes observed. Furthermore, the appearance of a recombinant rDNA molecule which has an intact Domain 2 characteristic of wild-type C3 rDNA in a mass culture of C3-<i>rmm4</i>/B-<i>pmr</i> heterozygotes confirmed that both 1-bp and 42-bp deletion confer a replication disadvantage. The significance of the emergence of recombination molecules is that it involved a crossover within the 25 bp that separates these two deletions (Figure 4A). This highly specific recombinational event precludes any doubt that the 1-bp and 42-bp deletions are the molecular determinants for the deficiencies on their respective rDNAs.

No recombination event was detected in the rDNA of Pmr sensitive cells from C3-<i>rmm1</i>/B-<i>pmr</i> mass cultures. DNA sequencing showed that the C3 rDNA still carried the 1-bp deletion associated with <i>rmm1</i>. One possible explanation for this observation is the failure, in rare cases, of the B-<i>pmr</i> rDNA to mature in the developing macronuclei at conjugation; such a failure would give rise to cells having exclusively C3-<i>rmm1</i> rDNA, which presumably would be maintained at normal copy number. An analogous observation has been made in wild-type C3/B heterozygous cultures, in which clones maintaining exclusively B rDNA have been described (Løvlie et al. 1988). An alternative explanation for the persistence of C3-<i>rmm1</i> in rare cells is the existence of a suppressor mutation that overcomes the replicative disadvantage of C3-<i>rmm1</i> rDNA. This or other more complex explanations have not yet been ruled out.

The recombinant rDNA molecules described here are homopalindromic; that is, the two halves of the palindrome are identical. One possible pathway for the production of recombinant homopalindromes is illustrated in Figure 4B. In the process of the homopalindrome formation,
the first cross occurred within the 25 bp region that separates the 1-bp and the 42-bp deletions leading to the recombinant intermediate, a heteropalindrome, which has been detected in Southern blots (data not shown). These intermediates rapidly recombined with each other to form the homopalindromes, which exhibited the best replication ability.

The production of molecules with a replication advantage through a recombination event observed in the mass cultures draws attention to the role of the *Tetrahymena* macronucleus in promoting somatic evolutionary adaptation. Infrequent macronuclear DNA recombination, which is coupled with the high ploidy of the macronucleus, provides the potential for a tremendous degree of somatic genetic diversity, and generates beneficial molecules that are available for natural selection during asexual reproduction. When circular rDNA plasmids were transformed into *Tetrahymena thermophila*, homologous recombination with endogenous macronucleus rDNA occurred, resulting in molecules larger than 21 kb that contained multiple copies of 5'NTS (Yu and Blackburn 1989; Yu and Blackburn 1990; Yu et al. 1988). Homologous somatic recombination phenomena have been implicated in a variety of molecular phenomena in eukaryotes, including DNA amplification (Schimke 1982), tumorigenic expression of oncogenes (Koufos et al. 1985), homozygosis of heterozygous loci (Potter et al. 1987), and homologous integration of transfected genes (Smithies et al. 1985).

The mutation in C3-rmm3 was located in the Type I repeat within the rRNA promoter. This result implies that rDNA replication and transcription could share a common regulatory factor, as reported in several DNA viruses in mammalian cells (see review in Heintz 1992; DePamphilis 1988). There is also evidence for interdependence of DNA replication and transcriptional regulatory sequences in non-viral systems. In *Drosophila*, amplification of the developmentally-regulated chorion genes requires an upstream region that also regulates transcription of chorion genes (Orr-Weaver and Spradling 1986). DNA replication and transcription can share cis-acting elements (Ariga et al. 1989; Brand et al. 1987; Iguchi-Ariga et al. 1988; Rivier and Rine 1992) or trans-acting factors (Buchman et al. 1988a; Buchman et al. 1988b).
Two DNA-binding proteins (GRFI and ABFI) from fission yeast can bind both replication ARS-associated elements and transcriptional regulatory elements (silencer and upstream activating sequences) (Buchman et al. 1988a; Buchman et al. 1988b). Moreover, the fact that several DNA-binding factors from a cellular extract of *T. thermophila* were shown to interact with Type I repeat elements (either the replication origin region or from the transcription promoter region) (Umthun et al. 1993), further supports the notion that DNA replication and transcription are highly interdependent.

Three independently-derived mutants (C3-nnm1, C3-nnm3, and C3-nnm4) exhibiting the same effect on replication verify the significance of the conserved Type I repeat in rDNA replication. The fact that we found mutations that altered rDNA replication across a region greater than 600 bp suggests that the origin is much larger than a typical prokaryotic origin of replication. Together these data provide genetic evidence that the Type I repeat is a *cis*-acting control element for DNA replication *in vivo*. The only other specific sequences implicated in the control of DNA replication in eukaryotes are ARS sequences in yeast (Campbell 1986) and amplification control elements (ACEs) in *Drosophila* chorion genes (de Cicco and Spradling 1984; Spradling and Mahowald 1981).
Figure 1. (A) Diagram of one-half of the rDNA palindrome, with the center at the left and telomere at the right (HH). In the enlarged 5'NTS region, arrow heads designating nuclease hypersensitive regions are redrawn from Palen and Cech (1984). Domain I and Domain 2, which correspond to the nuclease hypersensitive regions, and the origin of replication (ori) (Cech and Brehm, 1981) are indicated. (B) Organization of the repeated elements in the 5'NTS of rDNA alleles and summary of the positions of the mutations. Type I repeat, □, Type II repeat, □, and Type III, □, are the conserved repeated sequence elements found among several ciliate species.
Figure 2. Locations of the restriction site polymorphisms relevant to this study. The top row illustrates one-half of the rDNA palindrome. Domain 1 (D1), Domain 2 (D2) (shaded boxes), and the transcribed region (open box) are indicated. The lower two rows show relevant restriction sites BamHI (B) and SphI (S) on the C3 (middle row) and B rDNA (bottom row). Paromomycin resistant marker and the 42 bp deletion are designated by Pmr and Δ respectively.
Figure 3. DNA sequencing analyses of revertant rDNA molecules from C3-rmm1/B-Pmr and C3-rmm4/B-Pmr cultures. (A) Three independent XbaI-TaqI clones derived from C3-rmm4/B-Pmr show 11 As (wild-type C3 allele sequence) in the Type I repeat element (vertical hatched boxes). (B) Three independent XbaI-TaqI clones from C3-rmm1/B-Pmr show 10 As in the Type I repeat element (retaining rmm mutant characteristics).
Figure 4. (A) Diagram of crossover region between C3-*mm4* and B-Pmr producing a wild-type C3 rDNA sequence within Domain 2. "p" in the coding region designates the location of Pmr mutation. In the lower panel, the shaded area (25 bp range) represents the expected crossover location for production of recombinant rDNA. (B) Possible recombination pathway leading to the formation of rDNA devoid of the 1-bp deletion associated with C3-*mm4* and the 42-bp deletion associated with B rDNA. The relative replication rate is also indicated next to each molecule.
REFERENCES


PAPER II

ANALYSIS OF CONSERVED FEATURES IN THE REPLICATION ORIGIN REGION OF TETRAHYMENA rDNA
ANALYSIS OF CONSERVED FEATURES IN THE REPLICATION ORIGIN REGION OF TETRAHYMENA rDNA

Wen-Ling Shaiu

Drena D. Larson

Department of Zoology and Genetics

Iowa State University

Ames, IA 50011
ABSTRACT

An origin of replication, mapped by electron microscopy, is located in the 5' nontranscribed spacer (5'NTS) of ribosomal RNA genes (rDNA) of *Tetrahymena thermophila* (Cech and Brehm 1981). In order to understand the molecular mechanisms that control the regulation of rDNA replication in *Tetrahymena*, identification of the DNA sequence components associated with the origin region of DNA replication is essential. In the present report, the 5'NTS of C3 rDNA was analyzed for the presence of modular sequence elements which are shared among several putative eukaryotic chromosomal origins of replication (Benbow *et al.* 1992). These motifs include DNA unwinding elements, pyrimidine tracts, scaffold-associated regions, autonomously replicating sequences, transcriptional regulatory sequences, and, possibly, initiator binding sites. Sequence analyses indicate that several of these elements are clustered within the origin region in the 5'NTS of rDNA in *Tetrahymena*. Intrinsic DNA bending and scaffold-associated regions were identified and mapped to the origin region of rDNA. Neutral/neutral two-dimensional gel electrophoresis was employed to characterize replicative intermediates and to map sites of initiation in the rDNA. The results obtained indicated that replication initiation occurred within the 5'NTS of rDNA. Together with the previous finding that mutations separated by more than 600 bp can cause rDNA replication defects, these observations argue that the origin of *Tetrahymena* rDNA may encompass an unexpectedly large region. Furthermore, the organization of modular sequence elements and the mode of replication may be conserved among replication origins in eukaryotes.
INTRODUCTION

Investigation of the mechanisms that regulate eukaryotic DNA replication has been a major task in the field of cell biology for the last two decades. Little is known about the DNA sequences that are involved in initiation of DNA replication, except in prokaryotic and mammalian viral DNA origins, in which replication initiates at specific DNA sequences called replication origins (Bramhill and Kornberg 1988; Campbell 1986; Stillman 1989).

In contrast to the well-characterized prokaryotic and viral origins, eukaryotic chromosomal origins of replication have been difficult to isolate and characterize. The only well characterized specific sequence involved in eukaryotic DNA replication is the autonomously replicating sequence (ARS) in yeast. Some, but not all, ARSs in budding yeast, *Saccharomyces cerevisiae*, can function as chromosomal origins of replication (Brewer and Fangman 1991; Deshpande and Newlon 1992; Ferguson and Fangman 1992; Huberman et al. 1988; Walker et al. 1991; Zhu et al. 1992b). All ARSs share several features; first, they are high in AT content; second, an 11 bp ARS core consensus sequence (ACS) is indispensable for the function of the origin; and third, the flanking regions facilitate, but are not essential for, the ARS function (Campbell and Newlon 1991; Marahrens and Stillman 1992; Newlon 1988; Williams et al. 1988).

The origin of replication of rDNA of *Tetrahymena thermophila* was mapped by electron microscopy (EM) to the 5'NTS (Cech and Brehm 1981). Interestingly, two restriction fragments derived from the 5'NTS, and one fragment derived from the 3' nontranscribed spacer sequence (3'NTS), were shown to have ARS activity in yeast (Amin and Pearlman 1985; Amin and Pearlman 1986). One of these ARS activities co-localizes with the EM-mapped origin of replication. This region has been shown to be nucleosome-free and nuclease hypersensitive (Bonven and Westergaard 1982; Palen and Cech 1984), a common feature in viral and
prokaryotic replication origins, as well as in the yeast ARS1 origin (Brown et al. 1991; Umek and Kowalski 1988).

For some time, the eukaryotic DNA replication paradigm resembled the prokaryotic DNA replication model, in which replication starts at a specific initiation site, and DNA unwinding is coupled with DNA synthesis, to form a structure called the replication fork. Recently, several reports demonstrated that initiation events occur throughout broad zones in eukaryotic genomes, rather than at a specific origin sequences (Caddle and Calos 1992; Dijkwel and Hamlin 1992; Shinomiya and Sawako 1991; Vassilev and Johnson 1990; Vaughn et al. 1990). Sequence analyses of several origin regions in eukaryotes reveal that these origins share modular sequence elements (Benbow et al. 1992). Among the common motifs are DNA unwinding elements (DUEs) (having a lower free energy for unwinding than flanking regions) (Kowalski and Eddy 1989), pyrimidine tracts (that may serve as strong DNA polymerase-primase start sites) (Kaiserman et al. 1990), scaffold-associated regions (SARs) (that anchor the chromosome to the nuclear matrix) (Gasser et al. 1989), transcriptional regulatory sequences, and initiator binding sites.

Two features of replication origins that have received particular attention are bent DNA and SARs. Bent DNA structures usually are associated with a continuous run of A residues (2-6 bases) that occur with a 10-11 bp periodicity (Anderson 1986; Koo et al. 1986; Trifonov and Sussman 1980). Bent DNA sequences are implicated in both prokaryotic and eukaryotic origins of DNA replication (Anderson 1986; Caddle et al. 1990; Hernández et al. 1993; Lilley 1986; Linial and Shlomai 1988; Ryder et al. 1986; Snyder et al. 1986; Zahn and Blattner 1987). Furthermore, bent DNA sequences also occur in ARS elements that are isolated from a variety of organisms, including Drosophila (Gragerov et al. 1988; Mills et al. 1986), mammals (Bergemann and Johnson 1992) and plants (Eckdahl and Anderson 1989). In yeast, bent DNA has been reported to enhance ARS1 function on plasmids (Williams et al. 1988) although it is not essential for ARS1 function (Celniker et al. 1984; Marahrens and Stillman 1992).
Eukaryotic chromosomes are thought to be organized into a higher-order structure consisting of discrete and topologically-independent loop domains, which are fastened at their bases to the intranuclear framework by nonhistone proteins (Gasser and Laemmli 1987; Mirkovitch et al. 1984). Specific DNA sequences that exist at the bases of the DNA loops are called MARs or SARs (for matrix-associated regions, or scaffold-associated regions, depending on the isolation methods), and have been identified in a number of different species (see Phivan and Strätling 1990 and references therein). It has been proposed for many years that origins of replication may be associated with a nuclear structure in eukaryotic cells (van der Velden and Wanka 1987). For example, the fragment containing replication origin of α-globin genes in chicken is enriched in a fraction that was isolated from the high-salt matrix preparation (Kalandadze et al. 1990; Razin et al. 1986). Moreover, several putative DNA replication enzymes are enriched in nuclear matrix preparations (Gasser et al. 1989; Nakayasu and Berezney 1989). Fluorescence microscopy has been used to observe DNA replication at discrete sites in intact nuclei, and it has been shown that these sites are also the positions of DNA synthesis in nuclear matrix preparations (Mills et al. 1990; Nakayasu and Berezney 1989).

In this study, we characterized the structure and DNA sequence components of the replication origin in Tetrahymena rDNA, using a combination of computational sequence analyses and neutral/neutral two-dimensional (2-D) gel electrophoresis.
MATERIALS AND METHODS

Computational analyses of Tetrahymena rDNA sequences. The complete sequence of Tetrahymena rDNA from the GenEMBL database (accession number X54512) was analyzed using the Wisconsin Genetics Computer Group (GCG) Sequence Analysis Software Package, and the National Biosciences "Oligo" program. ENDS ratio plots were generated by the GentlBen program kindly provided by Dr. J. N. Anderson (VanWye et al. 1991).

Cultures and rDNA isolation. C3.V cells were grown in 2 % PPYS at 30° C until mid-log phase (about 2.5 x 10^5 cells/ml). Cells were starved in 10 mM Tris-HCl (pH 7.5) for 20 hours, and refed with 10 % PPYS and bromodeoxyuridine (BrdU, Sigma) to a final concentration of 1 % PPYS and 50 μg/ml BrdU for the indicated times. Total cellular DNA was isolated, and a 5 ml CsCl-Hoechst 33258 (Sigma) gradient was used to purify rDNA from the bulk nuclear DNA, as previously described (Spangler and Blackburn 1985).

Bent DNA analysis. A plasmid clone, pUC8-1, of the 5’NTS sequences from B rDNA was digested with Dral and/or XbaI. Restriction fragments purified using Genecon (Bio 101) were filled-in using Klenow polymerase (BRL), and ligated to a filled-in Xba I site of pBend2. Plasmid pBend2, kindly provided by Dr. S. Adhya (Kim et al. 1989), contains a circularly-permuted multiple cloning region. Restriction fragments that contain the inserts were separated on both 1.5 % agarose gels and 6 % polyacrylamide gels.

Isolation of scaffold-associated DNA. Tetrahymena macronuclei were isolated in a sucrose gradient according to the procedures of T. Higashinagakawa et al. (Higashinakagawa et al. 1992), and nuclear matrix isolation was performed essentially according to Dijkwel et al. (Dijkwel et al. 1991). Briefly, pellets of Tetrahymena cells were resuspended in 0.2 M sucrose/0.25 mM MgCl₂, and lysed with 0.2 % NP-40. Macronuclei were pelleted from a 2.1 M sucrose gradient in a Beckman Ti 70.1 rotor at 40,000 x g for > 2 hours. The pellets, which contained mainly the macronuclei, were washed with cell washing buffer (5 mM Tris-HCl [pH
7.4], 50 mM KCl, 0.5 mM spermine, 0.125 mM spermidine, 0.5% thiodiglycol, 0.25 mM phenylmethylsulfonyl fluoride [PMSF]), and incubated with stabilization buffer (5 mM Tris-HCl [pH 7.4], 50 mM KCl, 0.625 mM CuSO4, 0.5 mM spermine, 0.125 mM spermidine, 0.5% thiodiglycol, 0.25 mM PMSF, 0.1% NP-40) at 37°C for 20 minutes. Lithium diiodosalicylate (LIS) (Sigma) extraction buffer (10 mM LIS, 100 mM LiOAc, 0.1% NP-40, 0.05 mM spermine, 0.125 mM spermidine, 0.25 mM PMSF, 20 mM Heps-KOH [pH 7.4]) was used to lyse macronuclei. Treated macronuclei were pelleted in a table top low speed centrifuge and washed with matrix washing buffer (20 mM KCl, 70 mM NaCl, 10 mM MgCl2, 20 mM Tris-HCl [pH 7.4])/ 0.1% NP-40, followed by washing twice in 1X Hind III restriction buffer (10 mM Tris-HCl [pH 7.9], 50 mM NaCl, 10 mM MgCl2, 1 mM dithiothreitol). Hind III restriction digestion, using 1000 units, was carried out for greater than 6 hours at 37°C. The restriction reaction was spun down, and the supernatant was collected (the released DNA fragments were called loop DNA). Both matrix-associated DNA (MAD; DNA segments that remain attached to the nuclear matrix after Hind III digestion) and loop DNA were resuspended in TE/0.1% SDS with 200 μg of proteinase K and incubated at 50°C for at least 4 hours. After phenol extraction and ethanol precipitation, samples were subjected to restriction digestion, or stored in ethanol at -20°C.

**Two-dimensional gel electrophoresis.** The procedures for two-dimensional gel electrophoresis were based on those of Brewer and Fangman (Brewer and Fangman 1987). The replicative intermediates of rDNA was fractionated in 0.4% agarose gels in 1X TBE buffer (50 mM Tris-HCl [pH 8.0], 50 mM boric acid, 1 mM EDTA) at 1 volt/cm for 40 hours after restriction enzyme digestions. Gel strips containing the fragments of interest were cast into the top of a 1% secondary agarose gel. Secondary electrophoresis was carried out in 1X TBE buffer containing 0.3 μg/ml of ethidium bromide at 5 volts/cm for 5-8 hours.

Southern blots were performed using an alkaline transfer method (Reed and Mann 1985) to MagnaNT filters (MSI), and the probe used was the Type I repeat oligonucleotide (A-rich
strand 33 mers) labeled with $\gamma^{32}$P-dATP by T4 kinase (New England Biolabs). The probed filters were exposed to Kodak X-OMAT films with an intensifying screen at $-70^\circ$ C for one week.
RESULTS

Common modular sequence elements clustered within the NTSs. Computational analyses of the rDNA sequence from Tetrahymena revealed clustered common modular sequence elements that are shared with other eukaryotic origins of DNA replication. Figure 1 shows the distribution of these elements in one-half of the 21 kb rDNA molecule. Two major DUEs, which are inherently unstable duplex DNA segments characterized by their hypersensitivity to single-strand-specific nucleases (Umek and Kowalski 1990), are found in the 5'NTS of Tetrahymena rDNA; one major DUE is located in the 3'NTS. Interestingly, the major DUE distal to the palindromic center coincides roughly with the EM-mapped origin.

The sequence (Y)\textit{9}C(T)\textit{3}(Y)\textit{4} (pyrimidine tract) is a strong start site for DNA polymerase-primase, while (Y)\textit{7}R(C)\textit{3}(Y)\textit{5} is a weaker site (where Y=C or T; R=A or G) (Kaiserman et al. 1990). There were 3 perfect pyrimidine tracts (≥12 adjacent pyrimidines) found in the 5'NTS, 3 in coding region, and 2 in 3'NTS.

SAR consensus sequences were originally identified in Drosophila (Gasser et al. 1989; Rao et al. 1990), and four such sequences were used in the analysis; AATAAAAYAAA, TTWTWTTWTT, WADAWAYAWW, TWWTDTWWWW (where Y= C or T; W= A or T; D = A, G, or T). SARs seem to spread out along both NTSs, but with higher frequency in one of the major DUEs in the 5'NTS.

ARSs shown in Figure 1 are ≥ 9/11 match to the yeast ARS consensus (5'A\textit{7}TTAT\textit{A}T\textit{C}TTT\textit{A}T\textit{3'}). Because of overlapping ARSs, not all them are indicated in the figure. There are 28 ARS consensus sequences on the upper strand of the 5'NTS, and 45 on the lower strand; the corresponding figures for the transcribed region are nine and 10, and for the 3'NTS, 24 and 28. Both NTSs have clustered modular sequence elements; however, there were more of these modular sequence elements associated with the 5'NTS than the 3'NTS.
An intrinsically bent DNA segment is located in the origin-containing fragment. The potential for DNA bending due to tracts of A_{2-6} can be predicted using programs that calculate an ENDS ratio. The term ENDS ratio is defined as the ratio of the contour length of a nucleotide segment along the axis to the shortest distance between the ends of the segment (Eckdahl and Anderson 1987; VanWye et al. 1991). Figure 2 shows the ENDS ratio plot (computed at a window width of 120 nucleotides, and at a window step of 10 nucleotides) of one half of the rDNA palindrome. Both NTSs have relative high ENDS ratios, and the highest ENDS ratio appears in the 5'NTS, predicting bent DNA sequence(s) in this region. A direct test for the presence of bent DNA structures involves electrophoresis of the DNA fragment under study on agarose and polyacrylamide gels. A fragment containing a bent DNA sequence migrates on a polyacrylamide gel with an apparent molecular weight greater than its actual molecular weight; no discrepancy in migration is observed in electrophoresis on an agarose gel.

Specific fragments corresponding to the 5'NTS of rDNA from plasmid clones have been shown to exhibit anomalous migration in polyacrylamide gels but the position of bent DNA segments have not been precisely mapped (Christiansen et al. 1987). It is possible to localize bent segments because bent sequence located near the end of a fragment show less retardation than when located at the center of the fragment (Kim et al. 1989). To test whether the origin-containing fragment contains a bent DNA domain, a 430 bp XbaI-XbaI fragment from the 5'NTS of Tetrahymena rDNA was cloned into pBend2, which contains a set of circularly-permuted cloning sites (Figure 3A). This allow digestions by different restriction enzymes present in the cloning sites to generate a set of segments of the same contour length. Each member of the set has the rDNA insert at a different position relative to the center of the fragment. A recombinant clone, pXX59, was sequenced through the multiple cloning region and across one XbaI site to determine the orientation of the insert. Plasmid pXX59 was digested with six different restriction enzymes. All the rDNA containing fragments generated by the restriction digestions migrated at the same position on a 1.2 % agarose gel (Figure 3C);
in contrast, they exhibited anomalous migration on a 6 % polyacrylamide gel (Figure 3B). The different mobilities of fragments of the same size on polyacrylamide gels are indicative of the existence of DNA bending.

The results obtained from pXX59 indicate a bent DNA structure in the XbaI-XbaI fragment, which corresponds to the EM-mapped origin of replication. To further identify the bending region, a sub-fragment of XbaI-XbaI was analyzed. The segment of the 200 bp XbaI-DraI sequences exhibited no such anomalous migration on the polyacrylamide gels (data not shown). These results suggest that the bent DNA sequence is located within 3' half of the XbaI-XbaI fragment (see Figure 2).

**A scaffold-associated region is located near the palindromic center.** The strategy for mapping the scaffold-associated DNA sequence was similar to the one used by Razin's group (Kalandadze et al. 1990; Razin et al. 1986). Briefly, restriction fragments of recombinant clone pCT9 covering 5'NTS of C3 rDNA of *Tetrahymena* were separated on 1 % agarose gels and transferred to nylon filters. Labeled matrix-associated DNA (MAD; DNA segments that remain attached to the nuclear scaffold after Hind III digestion, which leaves the 5'NTS intact) and labeled loop DNA (DNA segments released from the scaffold structure after Hind III digestion) were used as probes for hybridization (see Materials and Methods).

Three fragments of lengths 820 bp, 700 bp and 430 bp are detected in blots probed with labeled MAD (Figure 4B, lane 1) and labeled loop DNA (Figure 4B, lane 2), corresponding to the cloned 5'NTS rDNA fragments of pCT9 (Figure 4A). The 3.1 kb fragment detected in these lanes represents vector DNA. The 700 bp fragment in lane 1 had the strongest relative intensity, when the hybridization signal of each fragment was normalized to the fragment length. Therefore, the MAD preparation was enriched in the 700 bp XbaI-HindIII fragment. This fragment contains the 680 bp TaqI-XbaI fragment next to the palindromic center, and 20 bp of vector sequence (Figure 4A). This result indicates that the TaqI-XbaI fragment contains DNA sequences corresponding to MAR/SAR.
Analysis of replicative intermediates by two-dimensional gel electrophoresis. In order to test whether replication initiation events occur in either of the DUEs of the 5'NTS, a 2-D gel electrophoresis technique was employed to detect bubble-containing replication intermediates. This technique has been used to map replication origins in many organisms (Bénard and Pierron 1992; Brewer and Fangman 1991; Gahn and Schildkraut 1989; Liang et al. 1993; Martín-Parras et al. 1991; Schwartzman et al. 1990; Yang and Botchan 1990).

The basis of 2-D gel electrophoresis technique is that the first dimension separates molecules according to their mass (or size) by electrophoresis at very low voltage, while the second dimension separates molecules primarily based on their shapes (branched structures) by electrophoresis at high voltage. Therefore, a given fragment containing either a single replication fork (Y arc), a centered or off-centered bubble (bubble arc or asymmetric arc), or two approaching forks (double-Y arc) can be distinguished (Figure 5A).

Southern blot of a typical result from 2-D gel electrophoresis is shown in Figure 5C. Unexpectedly, a strong Y arc but no bubble arc pattern was detected in the 2-D gel of BstBI digestion of C3 rDNA. It seemed that very few or no initiation events occurred within the 5'NTS of C3 rDNA. However, a closer examination of the region from 4.4 kb to 8.8 kb on an autoradiogram that was exposed for four weeks (the same filter as shown in Figure 5C), revealed a faint bubble-arc ascending from the 4.4 kb position (Figure 5D). A partial BstBI digestion could generate a 4.4 kb fragment with 2 inverted copies of the 5'NTS. If DNA molecules were captured before the nascent strands were fully synthesized, the unwound region of DNA would be single-stranded, and is resistant to restriction enzyme digestion. Therefore, it is possible that bubble-containing molecules with partially single-stranded regions exist in the 4.4 kb fragment. In addition, an intense hybridization signal at a unique spot near the 4.4 kb position was detected in the 2-D gels (marked by an arrow in Figure 5C). A similar observation has been reported in the rDNA repeat of *Saccharomyces cerevisiae* (Brewer and Fangman
1988; Linskens and Huberman 1988). This intense spot has been referred to as a replication fork barrier (Brewer et al. 1992).
DISCUSSION

The origin of replication of *Tetrahymena* rDNA has been roughly mapped by electron microscopy to a region 650 bp ± 300 bp from the center of the 21 kb palindrome in the B strain (Cech and Brehm 1981). A number of techniques have been employed to elucidate the mechanisms of replication regulation of the C3 rDNA allele, which exhibits the ability to outreplicate B type rDNA in a vegetative culture of the C3/B heterozygotes (Larson *et al.* 1986). Yu and Blackburn demonstrated that the signal(s) for replicative dominance is located within the 5'NTS of C3 rDNA, using transformation assays by microinjection (Yu and Blackburn 1989; Yu and Blackburn 1990; Yu *et al.* 1988; Tondravi and Yao 1986). Significantly, we have found that mutations resulting in a replication disadvantage are spread across about 600 bp in the 5'NTS (see Paper I of this dissertation).

Modular sequence elements that are present in other origins of replication, including DUEs, ARS, SAR, and pyrimidine tracts, also occur frequently in the 5'NTS of *Tetrahymena* rDNA. A bent DNA structure was found in the 430 bp Xba I-Xba I fragment, which roughly corresponds to the EM-mapped origin region. Bent DNA sequences have been reported to be associated with the origins of replication in several systems, for instance, bacteriophage λ (Zahn and Blattner 1985; Zahn and Blattner 1987), SV40 (Deb *et al.* 1986; Ryder *et al.* 1986), adenovirus (Anderson 1986), yeast ARS (Marahrens and Stillman 1992; Snyder *et al.* 1986; Williams *et al.* 1988), the DHFR gene of CHO cells (Caddie *et al.* 1990), the human c-myc gene (Bergemann and Johnson 1992), and the rDNA of pea (Hernández *et al.* 1993). It seems that a bent DNA structure is conserved among origins of replication from prokaryotes to eukaryotes. The role of bent DNA in replication is unclear, except in yeast, where it functions as an replication enhancer (Williams *et al.* 1988). It has been suggested that bent DNA functions as a structural landmark for replication proteins or transcriptional proteins (Linial and Shlomai 1987; Williams *et al.* 1988), e.g., RIP 60 and AP-1 recognize the sequences embedded...
in the bent DNA sequence of the DHFR origin region (Caddle et al. 1990). Alternatively, bent DNA could foster the interaction of proteins bound at distal sites (Stenzel et al. 1987).

Two fragments (TaqI-XbaI and XbaI-XbaI) from the 5'NTS and one (HindIII-EcoRI) from the 3'NTS of rDNA can replicate autonomously in yeast when cloned into plasmids (Amin and Pearlman 1985). Interestingly, all three ARS-containing fragments have ENDS ratios greater than 1.4 implying the existence of bent DNA sequences (see Figure 2). This strongly argues that bent DNA domain could enhance the ARS origin function, although, it might not function as a chromosomal origin in *Tetrahymena*.

DNA replication is thought to occur at the nuclear substructure, the nuclear scaffold (matrix), where SARs (MARs) are attached (Mills et al. 1990; Nakayasu and Berezney 1989; van der Velden et al. 1984). Some matrix-bound DNA fragments from human (Aguinaga et al. 1987; Sykes et al. 1988) and *Drosophila* (Amanti and Gasser 1990) were found to have ARS activities in yeasts. The *Tetrahymena* rDNA fragment that was shown to associate with the nuclear scaffold (Figure 4B) also functions as an ARS element in yeast (Amin and Pearlman 1986). The co-localization of the bent DNA structure and SAR motif of rDNA to ARS activity-containing fragments strongly supports the notion that the 5'NTS contains the potential origin of replication. Moreover, the organization of structural and functional domains within the replication origins seems to be conserved among eukaryotes.

Attempts to map the *in vivo* replication origin of rDNA using 2-D gel electrophoresis revealed strong Y arc 2-D pattern and a distinct replication fork barrier from the 2.2 kb BstBI-BstBI fragment (Figure 5C). However, bubble-containing molecules with partially single-stranded region were detected in a 4.4 kb fragment that contains 2 inverted repeats of the 5'NTS. This fragment apparently resulted from BstBI partial digestion. Several possibilities could explain the failure to detect bubble arcs from the 2.2 kb BstBI-BstBI fragment. First, replication may start very near one end of the fragment, such that bubble-containing DNA would rapidly become Y-shaped while replication progressed bidirectionally. Linskens et al.
reported that 2-D gel analysis cannot detect an origin which is located 16%-23% of the fragment length from the ends of a restriction fragment (Linskens and Huberman 1989). Second, the replicative intermediates may have been lost or deformed during the isolation steps, as it has been reported that replicative intermediates are rather unstable in short linear molecules (Zannis-Hadjopoulos et al. 1981).

Similar 2-D results have been reported in the analysis of the rDNA of pea (Hernández et al. 1993), i.e., replication forks were found within the putative origin of replication but no initiation events were detected within the fragment studied. Pea rDNA isolated from S-phase synchronized cells and enriched for replicative intermediates by benzoylated naphthoylated DEAE-cellulose chromatography generated the same results (Hernández et al. 1993). The authors interpreted their data as a low frequency of initiation events as has been reported in rDNA of yeast (Hernández et al. 1993).

Replicative intermediates of *Tetrahymena* rDNA were also isolated from synchronized cultures. Nevertheless, the degree of synchronization of rDNA replication by processes of starvation and refeeding has not yet been carefully analyzed. About only 1% of rDNA is replicating at the time of rDNA isolation (data not shown), and perhaps less than 2% of these intermediates are expected to contain replication bubbles smaller than 2 kb (see fig 4 in Cech and Brehm 1981). In other words, only about 0.02% of rDNA may contain bubbles small enough to be confined to the 5'NTS. Therefore, a highly synchronized culture or further enrichment of rDNA replicative intermediates may be required to allow detection of bubble arcs.

When the 2-D gel technique was applied to the DHFR gene of CHO cells, a broad zone of replication initiation was observed (Dijkwel and Hamlin 1992; Vaughn et al. 1990). By contrast, Okazaki fragment polarity analysis yielded data suggesting that the origin of DHFR gene is contained within a stretch of 450 bp or less (Burhans et al. 1990). In *Drosophila*, 2-D gel analyses of amplifying chorion genes on chromosome III revealed an amplification origin about 1 kb in size (Delidakis and Kafatos 1989), and an initiation zone of 4-8 kb (Heck and
Spradling 1990). However, replication of single-copy and tandemly repeated histone genes in early *Drosophila* embryo produced 2-D patterns suggesting random initiation throughout the studied regions (Shinomiya and Sawako 1991). In addition, data from the *ura4* gene in *S. pombe* suggest that the origin of replication can not be precisely located by 2-D gel analysis (Zhu et al. 1992a). Therefore, interpretations of 2-D gel patterns should be considered with caution (Umek et al. 1989; Zhu et al. 1992a).

2-D gel analyses from both pea and yeast rDNAs revealed the arrest of replication forks in the NTSs (Brewer and Fangman 1988; Hernández et al. 1993; Linskens and Huberman 1988). A similar result was obtained in *Tetrahymena* rDNA. The nature of the replication fork barrier is still unknown. However, it has been suggested that the replication fork barrier observed in yeast rDNA is not a consequence of collision between the replication machinery and RNA polymerase traveling in opposite directions (Brewer et al. 1992; Kobayashi et al. 1992), nor a consequence of the DNA structure itself in the pea rDNA (Hernández et al. 1993). DNA-protein complexes have been shown to be the cause of fork arrest in *E. coli*, yeast centromeres, and the EBV family of repeats (Gahn and Schildkraut 1989; Greenfeder and Newlon 1992; Kuempel et al. 1989).

In the case of *Tetrahymena* rDNA, the replication fork barrier probably does not come from the collision of DNA machinery and RNA polymerase, because the movement of the replication fork and the transcription process of rDNA are in the same direction, provided replication starts from the 5'NTS. DNA structure could possibly contribute to the fork arrest, but that remains to be tested. Several trans-acting factors recognizing sequences in the 5'NTS have been reported in *Tetrahymena*. One of these factors, the Type I-repeat specific binding protein, has been suggested to play a dual role in replication and transcription (Umthun et al. 1993). It is possible that this factor could contribute to the fork arrest, although more investigation is needed to understand the role of fork arrest.
The main conclusion of this report is that organization of *Tetrahymena* rDNA origin of replication is similar to the other eukaryotic origins and the rDNA origin may encompass a larger than expected region. To further define sites at which the initiation of rDNA replication occurs, techniques that provide higher resolution are required. Toward this end, a powerful technique, atomic force microscopy (AFM), will be employed to map the sites of initiation. AFM is routinely used to image DNA molecules with several advantages, e.g., fast preparation, high resolution (theoretically, atomic resolution), and the ability to image reactions in real time (Henderson 1992; Shaiu *et al.* 1993; Vesenka *et al.* 1992). It is possible to image tagged-DNA molecules and gain resolution of a few nucleotides. A strategy that applies the developed procedures (Shaiu *et al.* 1993), which involve AFM imaging of gold-labeled DNA molecules to detect the first nucleotides incorporated into replicating rDNA with bromodeoxyuridine will be used in the future.
Figure 1. The organization of modular sequence elements in one-half of the rDNA palindrome. Modular sequence elements depicted by symbols are described in the bottom of the figure. DNA unwinding elements with overall Tm 10-20°C lower than average for the sequence shown are classified as major DUEs. Minor DUEs have overall Tm 7-10°C lower than average. All DUEs longer than 100 bp are drawn. There are many single-mismatch SARs and pyrimidine tracts present, but only perfect matches to the SAR consensus sequences, and perfect pyrimidine tracts of at least 12 adjacent pyrimidine are indicated. Matches of sequences greater than 9 out of 11 bp ARS consensus sequence are shown. XbaI and TaqI sites are also indicated.
Major DNA Unwinding Element

Minor DNA Unwinding Element

Type I repeats

ARS Consensus (≥ 9/11 match)
(Not all ARS consensus sequences are shown)

Pyrimidine Tracts (≥12)

Scaffold Associated Regions
Figure 2. ENDS ratio plot. rDNA sequence was analyzed using the GentlBen program (120 bp window, 10 bp step) to generate ENDS ratio (middle panel). The percentage of A+T content is also graphed (bottom panel). A map of one-half of rDNA is shown in the top panel to depict the NTSs and transcribed region. EM mapped origin is designated as ori (Cech and Brehm, 1981). The solid boxes (■) and the ticks at the end (■■■) represent the Type I repeats and telomere repeats, respectively. Restriction sites: X, XbaI; D, DraI.
Figure 3. Bent DNA analysis. (A) Diagram of pBend2 with the circularly-permuted multiple cloning sites. XbaI and Sall are the only unique cloning sites available for bent DNA analysis. A schematic of insert fragments (open boxes) containing the bent sequence (shaded box) is shown at the right side of panel A. (B) The restriction digestion products were fractionated on a 6% polyacrylamide gel and a 1.2% agarose gel (C). Lanes 1 to 6 are digestion products of plasmid pXX59 by restriction enzymes BambII, XhoI, NruI, ClaI, PvuII, and MluI, respectively.
Figure 4. Southern analysis of scaffold-associated DNA. (A) Schematic of the 21 kb rDNA palindrome (top), and pT9 (bottom). The dashed line in the center and the ticks at both ends of the rDNA designate palindromic center and telomere repeats, respectively. D1 and D2 are nuclease hypersensitive domains as determined by Palen and Cech (1984). X, XbaI; H, HindIII; T, TaqI. (B) Plasmid pCT9 (TaqI-TaqI fragment from the 5'NTS of C3 rDNA inserted into pUC118) was digested with XbaI and HindIII. The digested products were electrophoresed on a 1% agarose gel, and transferred to nylon filters. Matrix-associated DNA (lane 1), and loop DNA (lane 2), were nick-translated with 32P-dATP, and used as probes of the digested plasmid fragments. The 0.7 kb XbaI-HindIII fragment is indicated by a solid arrowhead. Labeled 1 kb ladder was used as size marker (lane M).
Figure 5. 2-D gel analysis. (A) Diagrams of patterns generated by branched DNA molecules upon 2-D gel electrophoresis (see details in text) (Brewer and Fangman 1987). (B) A map of one-half of the Tetrahymena macronuclear rDNA palindrome. Restriction sites used in this study are indicated. Type I repeats are shown as small boxes in the 5'NTS. (C) Southern blot of a typical 2-D gel probed with the Type I repeat sequence. C3 rDNA was isolated from cultures that were starved and refed for 60 minutes before rDNA isolation. A big arrow indicates the position of accumulated replicative intermediates. (D) A longer exposure from the same filter is shown in panel D, in which a faint bubble arc (indicated by the small arrows) ascending from the 4.4 kb position.
REFERENCES


66


PAPER III

ATOMIC FORCE MICROSCOPY OF ORIENTED LINEAR DNA MOLECULES
LABELED WITH 5NM GOLD SPHERES
ATOMIC FORCE MICROSCOPY OF ORIENTED LINEAR DNA MOLECULES
LABELED WITH 5NM GOLD SPHERES

Wen-Ling Shaiu
Drena D. Larson
James Vesenka
Eric Henderson*

Department of Zoology and Genetics
Iowa State University
Ames, IA 50011

Keywords: AFM, streptavidin-gold, gene mapping, nanodissection
Running title: AFM of Gold Labeled DNA

*Corresponding author

Note: The source of this paper is Oxford University Press publication and is reproduced with permission of Oxford University Press.
ABSTRACT

The atomic force microscope (AFM) (Binnig et al. 1986) can image DNA and RNA in air and under solutions at resolution comparable to that obtained by electron microscopy (EM) (Bustamante et al. 1992; Hansma et al. 1992a; Hansma et al. 1992b; Henderson 1992a; Henderson 1992b; Yang et al. 1992). We have developed a method for depositing and imaging linear DNA molecules to which 5nm gold spheres have been attached. The gold spheres facilitate orientation of the DNA molecules on the mica surface to which they are adsorbed and are potentially useful as internal height standards and as high resolution gene or sequence specific tags. We show that by modulating their adhesion to the mica surface, the gold spheres can be moved with some degree of control with the scanning tip.
INTRODUCTION

The atomic force microscope (Binnig et al. 1986) is capable of generating high, sometimes atomic, resolution images of biological and non-biological surfaces (Drake et al. 1989; Hoh and Hansma 1992). DNA has been one of the most widely utilized test substances for imaging of biological samples by AFM (Bustamante et al. 1992; Hansma et al. 1992a; Hansma et al. 1992b; Henderson 1992a; Henderson 1992b; Vesenka et al. 1992a; Vesenka et al. 1992b; Yang et al. 1992). Several of these studies have shown that it is possible to immobilize and image DNA in air and under solutions at resolution comparable to that obtained by electron microscopy. Individual DNA molecules have been severed by the AFM at desired positions and small fragments removed (Hansma et al. 1992b; Henderson 1992b; Vesenka et al. 1992a). The ability of the AFM to image and dissect small pieces of DNA portends its utility as an instrument for gene isolation and manipulation.

We have begun to develop techniques for localization of specific sequences and directed recovery and manipulation of DNA fragments by AFM. Our approach has been to incorporate biotinylated nucleotides into DNA and subsequently react the biotinylated DNA molecules with a streptavidin-gold (SAG) conjugate, in this case 5nm diameter gold spheres. We show that DNA labeled with gold spheres can be deposited in an oriented fashion and reliably imaged by AFM. In addition, we demonstrate that the gold spheres can be moved easily, suggesting that they could be used as physical handles for manipulating attached DNA. Current work is focused on exploiting the properties of gold spheres to visualize specific genes in chromatin and to isolate the tagged DNA fragments.
MATERIALS AND METHODS

DNA labeling. Plasmid DNA, pUC119, was prepared by alkaline lysis and CsCl-EtBr gradient purification (Sambrook et al. 1989). DNA was 5' end-labeled with bio-dUTP (Enzo Biochem, NY) by using Klenow fragment of E. coli DNA polymerase I to fill in 5' overhanging ends generated by digestion with Hind III. Unincorporated bio-dUTP was removed by ethanol precipitation, and the bio-dUTP labeled DNA was resuspended in 10mM Tris-HCl (pH 7.2), 5mM MgOAc, 50mM NH4OAc, 1mM EDTA (TMNE) and incubated with 1 ml streptavidin-gold conjugate (Amersham) for 60 minutes at 25°C. The DNA-bound gold particles were separated from unbound gold conjugate by chromatography through Bio-gel A-50 (Biorad) in 20mM Tris-HCl (pH 7.5), 100mM NaOAc (TN). Fractions containing DNA were pooled and ethanol precipitated.

Sample deposition and AFM imaging. Biotin-streptavidin-gold-DNA (BSG-DNA) in 20mM Tris-HCl (pH 7.2), 100mM NaOAc and 5mM MgCl2 (TNM) (W. Rees, personal communication) was deposited directly onto freshly cleaved mica (Ted Pella, Inc.) for imaging in the AFM, or further concentrated by ethanol precipitation. The BSG-DNA was allowed to adsorb for 5 minutes (much shorter times worked equally well and in some experiments only a few seconds adsorption was required). The mica was rinsed with 1 ml ddH2O or dragged across a 10mM NH4OAc solution ten times and immediately and exhaustively dried with N2 gas with the gas flow orthogonal to the surface. DNA prepared in this way was imaged either under propanol or in air at relative humidity <10% (Bustamante et al. 1992; Hansma et al. 1992a; Hansma et al. 1992b; Henderson 1992a; Henderson 1992b; Thundat et al. 1992a; Thundat et al. 1992b; Vesenka et al. 1992a; Vesenka et al. 1992b; Yang et al. 1992). For this report, all images were collected on a Nanoscope II or Nanoscope III (Digital Instruments, Inc., Santa Barbara, CA) using Si3N4 tips (Digital Instruments, Inc., Santa Barbara, CA). In height mode images, the gray scale represents the sample height with
lighter features being taller. In error signal mode (Putman et al. 1992) images, the error signal is displayed, rather than the true image height. All height measurements were made on height mode images.

**Data preparation.** For data presentation, image files were imported into the Macintosh program Image 1.43 (custom settings: 400x400 or 512x512, 2048 header, 16 bit signed, swap bites, calibrate off) and converted to PICT format for preparation of figures in Aldus Freehand 3.1. Statistical analysis of height and length measurements was carried out on Kaleidagraph graphics program (Synergy Software, Reading, PA).
RESULTS

Images of unlabeled double-stranded DNA molecules in dry air (Bustamante et al. 1992; Thundat et al. 1992a; Thundat et al. 1992b; Vesenka et al. 1992a; Vesenka et al. 1992b) or under propanol (Hansma et al. 1992a; Hansma et al. 1992b; Vesenka et al. 1992b) were routinely collected (Fig. 1). The widths of DNA were narrower in propanol than air, similar to results reported elsewhere (Hansma et al. 1992a; Hansma et al. 1992b; Vesenka et al. 1992b). Since the two methods provided approximately equivalent lengths and heights under our conditions, and the air preparation is simpler, all measurements reported in this study were made on images collected in dry air.

Figure 2 shows the general procedure used for end-labeling DNA with bio-dUTP and SAG (see Materials and Methods for details). To obtain sharp images with a low background (i.e., with very little unbound SAG and/or salt deposits), the indicated purification steps were essential. Unlike EM, in which many contaminants are transparent to the electron beam, small contaminants such as salt deposits can obscure images of DNA molecules obtained by AFM. Therefore, after removal of unbound SAG by column chromatography, the DNA sample was further desalted by ethanol precipitation, deposited onto the mica from low or volatile salt solutions and rinsed after deposition with the same salt solutions or distilled water. The overall efficiency of labeling was approximately 60% under the conditions described in Materials and Methods. A number of molecules were observed with the appearance of a protein (i.e., streptavidin) attached to an end, but no gold sphere (data not shown). Detachment of streptavidin from SAG has been frequently observed by EM (E. Henderson, unpublished observation).

End-labeled DNA preparations appeared as single or multiple strands with gold spheres at one or both ends, as expected (Fig. 3). Although most SAG-containing molecules were terminally labeled (~70%), some molecules had SAG attached within the DNA strand. This
is most likely due to incorporation of bio-dUTP at nicks, since the plasmid DNA preparation contained ~30% nicked circular DNA prior to restriction enzyme digestion (data not shown).

Strikingly, the gold labeled molecules in a given field were often oriented in the same general direction. This suggests that the gold particle adhered first to the mica substrate during the drying process (with N2 gas) and that the DNA flowed in the direction of the N2 gas before it absorbed to the surface. In other words, the DNA was immobilized by the gold sphere at one end and extended in a direction determined by liquid and N2 gas flow. Molecules apparently lacking gold at an end (although possibly having bound streptavidin) also showed preferred orientation in some cases. Thus, the DNA may align to some extent in the N2 stream even in the absence of a gold end label. However, unlabeled molecules with an apparently random orientation were common, whereas non-aligned gold labeled molecules were infrequently seen.

The apparent contour lengths of 121 DNA molecules were measured (Fig. 4; see Materials and Methods). These molecules had an average length of 889 ± 91nm. The predicted value for the plasmid used is 1074nm in the B-form and 804 in the A-form (Saenger 1989). Since these molecules were imaged under very dry conditions (<10% relative humidity), the correspondence between the measured contour length and that expected for A-form (dehydrated) DNA may reflect a conformational property of the DNA helix.

The average height of the DNA was 0.54 ± 0.12nm (n=68, Fig. 4). The expected height of duplex DNA is approximately 2.0nm. The 4-fold discrepancy between the measured height and the expected height is not due to inaccuracy in the calibration of the Z-piezo since it was calibrated using atomic steps of highly oriented pyrolytic graphite and was within 20% of the expected value (2.6 Å) in all cases. Two possible explanations include sample compression due to pressure from the scanning tip (Bustamante et al. 1992) and embedding of the DNA in residual buffer salts (J. Vesenka, unpublished).
The height of the SAG measured by AFM was 5.8 ± 1.5nm (n=108, Fig. 4). This heterogeneity precludes use of currently commercially available SAG as an internal height standard (see Discussion). Determination of true sample width from measured values requires detailed knowledge about the tip shape (Bustamante et al. 1992; Keller et al. 1992; Vesenka et al. 1992a). Since this information is lacking, width measurements are not presented.

Height measurements were dependent on the scan direction. This dependence is a consequence of bending of the AFM cantilever due to strong interaction between the tip and the sample (or hydration layers on both) and is discussed in detail elsewhere (Radmacher et al. 1992; Thundat et al. 1992a; Thundat et al. 1992b; Zenhausern et al. 1992). For the height measurements reported here, the scan angle was optimized by piezo rotation at each area imaged to give overlapping oscilloscope tracings in both scan directions. The scan angle was frequently between 70° and 120°. This rotation often gives the most accurate height measurements (Radmacher et al. 1992), J. Vesenka, unpublished.)

It was possible to move the SAG spheres with the scanning tip by altering the humidity of the imaging environment. Similar results have been obtained with unconjugated gold spheres (J. Vesenka, unpublished). Figure 5 shows two examples of this process. The quality of the DNA image decayed as the humidity was increased as reported previously (Bustamante et al. 1992; Thundat et al. 1992a; Thundat et al. 1992b; Vesenka et al. 1992a). Concomitant with this decay was an increase in the imaging force, presumably due to increased meniscus forces between the tip and the sample. In the first series (Fig 5A), the force was not adjusted to compensate for the increase in humidity and the DNA became severed in several places at higher humidities. Movement of the SAG is indicated by diagonal lines in the scan taken at highest humidity (~ 60% relative humidity). When the humidity was subsequently reduced, the area previously scanned was visible, with the DNA apparently highly fragmented and gold accumulation at the edges of the field. In the second
series (Fig. 5B), the applied vertical force was monitored constantly and maintained at the initial level (10-20 nN). Again the gold was released from the surface at \( \sim 60\% \) relative humidity and the DNA was fragmented. These results demonstrate that changes in humidity can be exploited to manipulate colloidal gold spheres and DNA, but that methods to avoid concomitant damage to DNA samples must be established.
DISCUSSION

We have developed a method for imaging gold-labeled DNA molecules by AFM. Gold spheres 5nm in diameter are attached to DNA molecules through a biotin-streptavidin linkage. In addition to permitting localization of regions of the DNA into which the biotinylated nucleotide has been incorporated, this method allows the oriented deposition of linear DNA molecules. Since the AFM and related microscopes have atomic resolution on hard surfaces, it has been suggested that they have considerable potential for use in direct DNA sequencing (Hansma et al. 1991). Methods for deposition of linear DNA fragments in an oriented and elongated fashion, such as that described here, would facilitate this effort. In the near term, since the method presented here is simple and rapid, it may be possible to develop techniques in which direct visualization of DNA samples with the AFM would complement or replace analytical gel electrophoretic methods (e.g., for restriction mapping) or electron microscopy methods (e.g., for mapping DNA replication origins). Accurate mapping of DNA locations by AFM requires that contour length measurements be precise. In this study, images were collected under conditions expected to facilitate the formation of A-form DNA. Under these conditions we found that the average contour length for the molecules measured was within 11% of the expected contour length for A-form DNA based on an average axial rise of 2.56Å/bp (Saenger 1989). Therefore, these results suggest that the AFM has the potential for use in mapping experiments within 11% error. This value can easily be improved by, for example, higher pixel density and more sophisticated contour length measuring systems. Moreover, it is possible that the average value used to calculate the expected A-form contour length for the plasmid used in this study may be misleading since it does not account for sequence specific variation in the axial rise value.

Since the AFM can be used to dissect as well as image DNA, physical handles such as gold spheres may be useful for manipulation and recovery of DNA fragments generated by
AFM scission. In this study, the SAG was manipulated readily with the scanning tip but only at the cost of fragmenting the attached DNA. Future efforts will focus on devising methods for recovering specific DNA molecules in intact form. Development of techniques for imaging and manipulating DNA under aqueous conditions will greatly facilitate this effort.

The use of 5nm gold spheres as tags could, in principle, provide a convenient size calibration standard. This type of non-compressible standard (J. Vesenka, unpublished) would be extremely useful in assessing the degree of compression of samples by the scanning tip. If the measured height of the gold spheres is not accurate due to piezo calibration error one could still calculate the true height of the sample (in its compressed form) by a simple scaling procedure using the known gold sphere diameter determined by another method (e.g., EM). To serve as a reliable scaling standard, however, the gold spheres must be homogeneous in size, a requirement clearly not met by the SAG used in this study. Commercially available unconjugated gold spheres that are much more homogeneous in size than the SAG used here function well at size standards when mixed with biological samples (J. Vesenka, unpublished). Since attachment of gold spheres to the sample is not required if they are to serve as height standards only, their use should dramatically improve the accuracy of the height measurements in the AFM.

Methods for single molecule manipulation and imaging by AFM are rapidly evolving. Based on the demonstrated capabilities of the AFM it is likely to become a powerful tool for gene manipulation and physical mapping, a major near term goal of the Human Genome Project.
ACKNOWLEDGMENTS

The authors are grateful to Daniel Jondle for help with height measurements, and to the reviewers for helpful comments and criticisms. This work was supported by NSF grant DIR-9004649 (EH), NIH grants GM41023 (EH) and GM41899 (DDL) and Iowa State University (EH and DDL). Journal Paper No. J-15095 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 3064.
Figure 1. Typical height mode (see Materials and Methods) images of plasmid DNA molecules collected under propanol and in dry air (relative humidity ~10%). In these gray scale images tall features are lighter and flat features are darker. The gray scale bar on the right indicates the gray level spectrum and corresponding feature heights.
Figure 2. Diagram of the procedures used for labeling, purification and deposition of DNA samples for AFM imaging. Details and buffer abbreviations are in Materials and Methods.
Figure 3. Typical field and gallery of individual end labeled DNA molecules. The labeled DNA molecules are oriented in the same general direction as a consequence of the drying procedure. White arrowheads indicate molecules with internal labels. The internal labeling results from nick translation of nicked circular plasmid DNA during the end labeling procedure. Black arrowhead indicates a multimeric structure commonly observed, presumably resulting from biotin/streptavidin network formation. Images were taken in error signal mode (see Materials and Methods). Bar = 500nm.
Figure 4. Histogram showing the contour length distribution for the DNA and the height distributions for the DNA and SAG used in this study.
Figure 5. Humidity regulated movement of gold spheres by the AFM scanning tip. (A) The applied vertical force increased to greater than 120 nN as the humidity increased, resulting in DNA damage prior to release from the substrate (white arrowheads). (B) The applied vertical force was maintained between 10-20 nN. DNA damage was still evident in this experiment. In both cases, the last frame shows a larger field scan of the area shown in the preceding frames (made after reducing the humidity to < 26%). The gold spheres have accumulated at one edge. Relative humidity is indicated in the corner of each frame. Images were taken in error signal mode. Bar = 250nm.
REFERENCES


PAPER IV

VISUALIZATION OF CIRCULAR DNA MOLECULES LABELED WITH COLLOIDAL GOLD SPHERES USING ATOMIC FORCE MICROSCOPY
VISUALIZATION OF CIRCULAR DNA MOLECULES LABELED WITH COLLOIDAL GOLD SPHERES USING ATOMIC FORCE MICROSCOPY

Wen-Ling Shaiu, James Vesenka, Daniel Jondle, Eric Henderson, and Drena D. Larson*

Department of Zoology and Genetics
Signal Transduction Training Group
Iowa State University
Ames, IA 50011

*Corresponding Author

Drena D. Larson
Department of Zoology and Genetics
2114 Molecular Biology Building
Iowa State University
Ames, IA 50011
(515)294-1112 voice
(515)294-2876 FAX
ddlarson@iastate.edu

Note: The source of this paper is J. Vac. Sci. Technol. and is reproduced with permission of American Institute of Physics.
ABSTRACT

We have imaged gold-labeled DNA molecules with the Atomic Force Microscope (AFM). Circular plasmid DNA was labeled at internal positions by nick-translation using biotinylated dUTP. For visualization, the biotinylated DNA was linked to streptavidin-coated colloidal gold spheres (nominally 5 nm diameter) prior to AFM imaging. Reproducible images of the labeled DNA were obtained both in dry air and under propanol. Height measurements of the DNA and colloidal gold made under both conditions are presented. The stability of the DNA-streptavidin colloidal gold complexes observed even under propanol suggests that this labeling procedure could be exploited to map regions of interest in chromosomal DNA.
INTRODUCTION

The potential capability of the AFM to generate atomic resolution images of biological surfaces has motivated immense interest in imaging of DNA molecules (Bustamante et al. 1992; Hansma et al. 1992a; Hansma et al. 1992b; Henderson 1992a; Henderson 1992b; Thundat et al. 1992a; Thundat et al. 1992b; Vesenka et al. 1992a; Vesenka et al. 1992b; Yang et al. 1992; Zenhausem et al. 1992). Currently, AFM images of DNA in air and under solutions routinely display resolution comparable to that obtained by electron microscopy. The AFM can also be used to manipulate DNA. Individual plasmid DNA molecules have been cut by the AFM scanning tip and the excised fragments moved (Bustamante et al. 1992; Hansma et al. 1992b; Henderson 1992b). The high resolution imaging and manipulation capabilities of the AFM portend its utility as an instrument for chromosome mapping, gene manipulation, and possibly DNA sequencing (Hansma et al. 1991).

We have investigated the possibility of imaging DNA tagged with modified nucleotides that could serve as site specific markers for mapping of chromosomes and, potentially, as physical handles for manipulating DNA. In previous work, a method for imaging linear DNA molecules with gold spheres attached specifically at their ends was developed (Shaiu et al. 1993). This was achieved by incorporating biotinylated nucleotides into DNA and reacting the biotin-containing DNA with a streptavidin-gold conjugate (SAG). In this report, we demonstrate that circular DNA molecules can be labeled with gold at internal positions and reproducibly imaged both in air and under propanol. Differences in the measured heights of DNA and gold spheres under different imaging conditions and the implications of these results for reliable imaging of biological samples in the AFM are discussed.
MATERIALS AND METHODS

DNA labeling. Supercoiled plasmid DNA (pUC119) was prepared by alkaline lysis and CsCl2-EtBr gradient purification. Purified DNA was nick-translated with biotin-dUTP (Enzo Biochem, NY) DNase I and *E. Coli* DNA polymerase I for 20 minutes at 37°C. Unincorporated biotin-dUTP was removed by ethanol precipitation, and the bio-dUTP labeled DNA was resuspended in 10mM Tris-HCl (pH 7.2), 5mM MgOAc, 50mM NH4OAc, 1mM EDTA (TMNE) and incubated with 2 ml streptavidin-gold conjugate (SAG; Amersham) for 60 minutes at 25°C. The DNA-bound gold particles were separated from unbound gold conjugate by chromatography through Bio-gel A-50 (Biorad) in 20mM Tris-HCl (pH 7.5), 100mM NaOAc (TN). Fractions containing DNA were pooled and ethanol precipitated.

Sample deposition and AFM imaging. Biotin-streptavidin-gold-DNA (BSG-DNA) in 20mM Tris-HCl (pH 7.2), 100mM NaOAc and 5mM MgCl2 (TNM) was deposited directly onto freshly cleaved mica (Ted Pella, Inc., Redding, CA) for imaging in the AFM, or further concentrated by ethanol precipitation. Samples were prepared for imaging as previously described (Shaiu et al. 1993). All images were collected on a Nanoscope II (Digital Instruments, Inc., Santa Barbara, CA) using Si3N4 tips (Nanoprobes). All images displayed were collected in error signal mode (Putman et al. 1992) (ESM). In ESM images the error signal generated as the piezo attempts to compensate for cantilever deflection is displayed, rather than the true sample height. In height mode (HM) images, a gray scale typically represents the sample height with lighter features being taller. All height and width measurements were made on height mode images. The Z-piezo was calibrated using step edges on graphite.
RESULTS AND DISCUSSION

We have previously reported a procedure for labeling linear DNA molecules specifically at the ends with biotin-streptavidin-gold (BSG) (Shaiu et al. 1993). In this study, a nick-translation protocol was used to incorporate label at random internal positions in circular plasmid DNA molecules (see Materials and Methods). Fig. 1 shows a typical field and gallery of internally labeled DNA molecules imaged in dry air (relative humidity < 10%). The non-random orientation of the molecules is the consequence of the gold labeling and drying procedure and has been discussed elsewhere (Shaiu et al. 1993). Most of the labeled molecules are circular, as expected, with a single gold sphere attached. Under nick-translation conditions used in this experiment, approximately 30% of the DNA molecules appear to contain BSG, although much higher labeling efficiencies were achieved by increasing the length of the labeling reaction. In such cases, however, the labeled DNA often contained so many gold spheres that it was difficult to visualize the DNA (data not shown). Despite the brief DNA labeling protocol used in this study, it was initially surprising to observe the majority of molecules bound by a single gold particle. However, upon closer scrutiny it was apparent that many of the DNA molecules had unconjugated streptavidin protein attached to them at multiple sites (Fig. 2, white arrowheads). EM studies have demonstrated that commercially available colloidal gold-protein conjugates often contain a large proportion of free protein (E. Henderson, unpublished results). Thus, the efficiency of streptavidin-biotin complex formation was probably significantly higher than that calculated based on the number of gold particles observed bound to DNA.

Unlike the EM, in which many contaminants are transparent, the AFM detects stable features it encounters on a surface with little prejudice. Therefore, AFM samples must be very clean. In this and the previous study (Shaiu et al. 1993), two purification steps were essential for obtaining reliable images of BSG-DNA with low background: 1) removal of
unincorporated streptavidin-gold conjugate by column chromatography and 2) removal of salts (components of the labeling reaction buffer) by ethanol precipitation (see Materials and Methods for details).

AFM imaging of DNA under propanol has been shown to give better resolution and narrower apparent widths than imaging in air (Hansma et al. 1992a; Hansma et al. 1992b; Vesenka et al. 1992a; Vesenka et al. 1992b). Fig. 2 shows a gallery of images collected from the same labeled DNA sample, imaged with the same tip in dry air (top row) and subsequently under propanol (middle row). In this case, differences in the apparent width are small, with the air images appearing slightly thinner. However, DNA imaged in air with different tips often appeared much thicker (see Fig. 1). While innate differences in individual tip geometry clearly affect the apparent DNA thickness, an explanation for observing this phenomenon more frequently in air than in propanol may be that the tip is more susceptible to becoming charged and picking up surface contaminants (e.g., salt residue) in air than in propanol where static electrical charges may be more easily dissipated and the tip is constantly immersed in a cleansing solvent. In addition to tip shape, other imaging parameters can alter the apparent width. For example, in error signal mode (Putman et al. 1992) the apparent width of the DNA is dependent upon the gain settings. Finally, the appearance of DNA in the AFM is critically dependent upon the sample preparation method. BSG-DNA samples imaged in dry air showed a higher frequency of circular DNA whereas the same sample imaged under propanol sometimes appeared to contain primarily linear molecules (Fig. 2, third row). This is probably due to the use of excessive air force during drying of the sample on the mica surface, since a higher proportion of circular molecules was observed under propanol when care was taken to dry the DNA sample under very gentle air flow (Fig. 2, middle row). Apparently, most of the molecules with a linear appearance under propanol are supercoiled and/or nicked circular molecules in which two halves of the circular DNA molecules have aggregated locally, or "collapsed" to form half-sized linear structures.
Others have observed similar effects of drying and spreading conditions on the appearance of circular DNA molecules in the AFM (T. Thundat, personal communication).

The height of DNA molecules and gold spheres was measured both in air and under propanol (Fig. 3). The average height of DNA was 1.11 ± 0.47 (n=101) measured under propanol and 0.59 ± 0.18 nm (n=148) measured in air. The average height of streptavidin-colloidal gold spheres was 3.9 ± 1.8 nm (n=109) measured under propanol and 5.1 ± 3.2 nm (n=238) measured in air. T-tests performed on these data indicated that the differences in heights measured in air vs propanol were statistically significant (for DNA, T=12.12, p=0.0000; for gold spheres, T= 3.75, p=0.0002). Although the apparent height of the gold under both imaging conditions was close to that expected, the average height in propanol was somewhat less than in air. The DNA showed a striking deviation from the expected height (~2 nm) under both conditions and had an apparent height in propanol nearly double that observed in air.

Bustamante and colleagues (Bustamante et al. 1992) have shown that increased scanning force leads to decreased apparent height in repulsive contact mode imaging (~100nN) in air. Interestingly, the measurements reported here were made in the attractive contact mode in air (with negative cantilever deflection, due to meniscus forces (Weisenhorn et al. 1989)) and at much lower forces (~10nN), yet the DNA heights measured were still less than expected. Several authors have shown that apparent sample height can vary as a function of humidity (Thundat et al. 1992b) and frictional effects (Zenhausern et al. 1992). Although we have optimized the scan orientation so that we only obtain positive contrast images (Shaiau et al. 1993), cantilever bending under frictional load could have contributed to the differences in apparent heights observed in air and in propanol. Different adhesive interactions between the tip and the sample in the two imaging environments may alter the frictional effect and/or result in physical compression of the sample, even in the attractive contact imaging mode. Finally,
salt residue surrounding the sample or accumulated on the tip due to electrostatic charges could play a role in the height differences observed.
CONCLUSIONS

This report demonstrates the capability of AFM to image DNA molecules tagged at internal positions with biotinylated nucleotides. The tagged positions were visualized by reacting them with a streptavidin-colloidal gold complex. A critical point is that these complexes were stable during imaging both in dry air and under propanol. This approach could be exploited for physical mapping of genes or regulatory regions within chromosomal DNA. For example, in EM studies of DNA replication, sites of nascent DNA synthesis have been observed by incorporation of biotinylated nucleotides into DNA in vivo, followed by visualization of the sites of incorporation using streptavidin-colloidal gold or streptavidin-ferritin conjugates (Hiriyana et al. 1988). Our results suggest that a similar protocol could be used to map replication origin regions using the AFM, with the advantages of much simpler sample preparation and potentially higher resolution. Moreover, since the AFM can distinguish unconjugated protein bound to DNA, satisfactory mapping results may be obtained using free protein rather than protein-colloidal gold conjugates. Finally, these results imply that the AFM may be a useful tool to investigate many types of DNA-protein interactions.

The use of incompressible gold spheres as internal height standards may provide a means to assess tip geometry and the degree of specimen damage by the scanning tip during data collection (J. Vesenka, unpublished results). Unfortunately, the gold conjugates currently available are very heterogeneous in size and therefore relatively poor height standards. If gold spheres are used as height standards only, homogeneous preparations of unconjugated gold that are commercially available are the method of choice. However, direct attachment of gold spheres (even heterogeneous preparations) may prove very useful as a physical mechanism for manipulating DNA molecules to which they are attached (Shaiu et al. 1993).
ACKNOWLEDGMENTS

We thank Audra Krekyes for assistance in height calibration, John Weber of help with data acquisition. We are especially grateful to Jamie Weiss for the T-test calculations. This work was supported by NIH grants GM41899 (D.D.L) and GM 41023 (E.H.) and NSF grants DIR-9004649 (E.H.) and DIR-9113593 (P. Haydon, P.I.).
Figure 1. Typical field and gallery of circular DNA molecules labeled at internal positions with biotin-streptavidin-gold (see Materials and Methods). All images were taken in the error signal mode (Putman et al. 1992) in dry air (< 10% relative humidity). Bar = 250nm.
Figure 2. Gallery of individual internally labeled DNA molecules. The top two panels show examples of DNA molecules from the same sample preparation, imaged first in dry air and subsequently under propanol, using the same scanning tip. The bottom panel shows images obtained from the same DNA sample imaged under propanol with a different scanning tip. The rod-like appearance of the circular DNA molecules in the bottom panel is probably due to harsher drying conditions used during spreading (see text). White arrows indicate positions in which unconjugated streptavidin protein appears to be bound to DNA. All images were taken in the error signal mode. Bar = 250nm.
Figure 3. Histograms showing the apparent height distributions for the DNA and SAG used in this study. Measurements made under propanol are shown in the panels on the left; those made in dry air are shown in the panels on the right. All measurements were taken from height mode images.
REFERENCES


GENERAL CONCLUSIONS

Eukaryotic chromosomal origins of replication are suggested to be larger and more complex than the prokaryotic ones. The goal of this research is to define the origin of replication in rDNA of *Tetrahymena*. Therefore, the experimental designs are directed at testing the above hypothesis. Together, studies in *Tetrahymena* rDNA support this hypothesis.

The studies of the C3-rmm mutants demonstrate that the conserved Type I repeat is a cis-acting control element involved in rDNA replication. In the studies reported here, three independent rmm mutants exhibited a replicative disadvantage in macronuclei when crossed with wild-type allele B. A rare somatic recombination event, which occurred in the long-term vegetative culture of C3-rmm4/B-Pmr heterozygotes, reverted the rmm4 allele with a replicative disadvantage into one displaying replicative dominance. This observation further supports the finding that the 1-bp and the 42-bp deletions that affect the integrity of the Type I repeat are the determinants for the replicative disadvantage phenotypes. The Type I repeat is, therefore, the only cis-acting element characterized in vivo thus far that affects the non-amplification mode of DNA replication in eukaryotic systems, apart from the ARSs in yeast.

Although C3-rmm3 rDNA exhibits replication disadvantage, its site of mutation is located in a Type I repeat in the transcription promoter region, about 100 bp upstream of the transcription start point. Together with the finding that at least one trans-acting factor specifically recognizes the Type I repeat, including the promoter Type I repeat (Umthun et al. 1993), analysis of C3-rmm3 lends support to the notion that regulation of transcription and replication are inter-related, and that both processes may share common cis- and/or trans-acting factors.

Computer sequence analyses, and 2-D gel electrophoresis assays of replicative intermediates, have suggested that the organization of the origin and the mode of replication in *Tetrahymena* rDNA are very similar to those of other eukaryotic origins of replication. It may
reasonably be concluded that the size of the *Tetrahymena* rDNA origin may encompass a larger region, as happen in many other eukaryotic origins.

A model for eukaryotic origins

From studies of the replication origin of *Tetrahymena* rDNA, and analyses from four eukaryotic chromosomal origins (Benbow et al. 1992), a model of a typical eukaryotic origin of replication is advanced. Several general features of this model origin are described. First, the origin is not dependent on the primary DNA sequence (yet it is high in A+T content), but on its relative free energy for unwinding, and on the presence of alternative DNA structures and conserved modular elements. Second, the size of the origin varies depending on the distribution of all the modular elements and the structural and functional domains. In other words, these cis-acting sequences can be clustered within a small segment, or scattered across a large region, thereby defining the origin. The smallest origin known in yeast (ARSs) is about 100-200 bp, and the largest could be the size of a eukaryotic replicon, approximately 50-100 kb.

The putative origins would have at least one DNA unwinding element (DUE), more than one pyrimidine tract, scaffold-associated regions, autonomously replicating sequences (ARSs), transcriptional regulatory sequences, initiator protein binding sites, and alternative DNA structures. The DUE is the region that exhibits low free energy for unwinding; therefore, it is the first region unwound. Pyrimidine tracts may serve as preferred start sites for the action of DNA polymerase-primase (Kaiserman et al. 1990). Scaffold-associated regions anchor DNA to the nuclear substructure, such that cells can coordinate the replication of large numbers of replicons (Gasser et al. 1989). Fortuitous matches to the ARS core consensus sequence may be because of the high AT content in the origin; nevertheless, the presence of such modular elements suggests the ability of autonomous replication of the region on plasmids. Transcriptional regulatory factors that interact with sequences within the origin region either foster the assembly of the replication complexes, or are involved in temporal coordination of
replication and transcription processes (Heintz 1992). Initiator binding sites may be redundant to the bent DNA sequence, since bent DNA is one of the alternative structures and has been suggested to serve as a recognition site for DNA replication initiation (Linial and Shlomai 1988; Williams et al. 1988).

Future directions

I have successfully imaged DNA molecules tagged in vitro with colloidal gold by atomic force microscopy (AFM). This technique can be a very powerful tool for gene manipulation and physical mapping. Therefore, the immediate future work will be mapping the first nucleotides incorporated into the nascent strands. It is feasible to locate the first nucleotides incorporated into the replicating rDNA, upon starvation and refeeding of Tetrahymena in the presence of bromodeoxyuridine (BrdU), by AFM. Nascent rDNA tagged in vivo in this manner may be detected with antibody-gold conjugates by AFM imaging, as described in Paper III and Paper IV. The sites of BrdU incorporation can be resolved at 1 to 5 nm (about 10 bp resolution), depending on the colloidal gold used.

A second direction will be constructing synthetic origins that contain the modular elements and the structural domains mentioned in the proposed model, and testing of these synthetic origins in vitro using a cell-free DNA replication system, or in vivo using transformation assays. A series of origins containing different combinations of modular elements, in addition to linker sequences that have no sequence relationship to the modular elements in the origin region, will be constructed. A Xenopus in vitro replication system offers an excellent milieu in which to study the replication ability of the synthetic origins in vitro, while Tetrahymena microinjection is a well characterized in vivo system for testing the synthetic origins. Techniques like AFM and 2-D gel electrophoresis will be used to examine the replicative intermediates generated by the input synthetic origins.
REFERENCES


ACKNOWLEDGMENTS

First, I would give my deep appreciation to Dr. Drena Larson for her guidance and effort in training me as a scientist. Second, I am grateful for the help of Dr. Eric Henderson and Dr. James Vesenka in AFM work. I would also like to acknowledge all the members in Larson's and Henderson's labs. Their assistance in the lab work has made my job easier.

I'd also like to express my sincere gratitude to James Caffrey, my husband, for the spiritual and practical support through my graduate career. In addition, Alexander (Bo-Bo) Caffrey, my son, has been very cooperative during the time of my writing. I am so grateful.