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.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9335033

**Characterization of DNA-binding proteins that recognize the
replication origin of the rRNA genes of *Tetrahymena thermophila***

Umthun, Angela Rae, Ph.D.

Iowa State University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Characterization of DNA-binding proteins that recognize the
replication origin of the rRNA genes of *Tetrahymena thermophila*

by

Angela Rae Umthun

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

Members of the Committee:

Signature was redacted for privacy.

Iowa State University
Ames, Iowa

1993

For my parents and my love, Steve

TABLE OF CONTENTS

LIST OF FIGURES	vi
ABSTRACT	viii
GENERAL INTRODUCTION	1
An explanation of the dissertation organization	8
PAPER 1. A SINGLE-STRANDED DNA-BINDING PROTEIN BINDS SPECIFICALLY TO A REPLICATION CONTROL ELEMENT IN THE ORIGIN REGION OF <i>TETRAHYMENA</i> <i>THERMOPHILA</i> rDNA	10
ABSTRACT	12
INTRODUCTION	13
MATERIALS AND METHODS	18
Cell growth and S100 extract preparation	18
Hydroxyl radical footprinting	18
Gel mobility shift assays	19
Column chromatography	20
RESULTS	22
Identification of DNA binding proteins that recognize Type I repeat sequences	22
Separation of duplex and single-stranded Type I binding factors	25
Interaction of the duplex Type I binding factor (ds-TIBF) with the AT-rich central region of the Type I repeat	30

Interaction of the single-stranded Type I binding factors ssT-TIBF and ssA-TIBF with T-rich and A-rich oligonucleotides	35
Sequence-specific interactions of ssA-TIBF with the Type I repeat	35
DISCUSSION	41
ds-TIBF	42
ssT-TIBF	43
ssA-TIBF	44
Sequence-specific single-stranded DNA-binding proteins	46
Defining the functional origin of replication of <i>Tetrahymena</i> rDNA	46
A possible role for the Type I repeats in both transcription and replication	48
ACKNOWLEDGMENTS	50
REFERENCES CITED	51

PAPER 2. A DNA-BINDING PROTEIN INTERACTS WITH BOTH THE REPLICATION ORIGIN AND THE TRANSCRIPTIONAL PROMOTER OF THE rRNA GENES IN <i>TETRAHYMENA</i> <i>THERMOPHILA</i>	64
ABSTRACT	66
INTRODUCTION	67
MATERIALS AND METHODS	73
Cell growth and S100 extract preparation	73
Column chromatography	73

Gel mobility shift assays	75
UV crosslinking	76
RESULTS	77
The DNA-binding activity of ssA-TIBF resides in a 95 kD polypeptide	77
ssA-TIBF binds to additional sequences on the 3' end of the Type I repeat	82
Comparison of sequences surrounding the Type I repeats in the 5' NTS	83
The binding affinity of ssA-TIBF for C3 and B extended Type I repeats mimics the difference in replication efficiencies between C3 and B rDNA <i>in vivo</i>	88
ssA-TIBF specifically interacts with the Type I repeat that functions as an essential element of the promoter <i>in vitro</i>	91
DISCUSSION	95
REFERENCES CITED	99
 GENERAL CONCLUSIONS	107
Model: Roles of ssA-TIBF, ssT-TIBF, and ds-TIBF in rDNA replication	116
Future direction	118
LITERATURE CITED	122
ACKNOWLEDGMENTS	143
APPENDIX: ds-TIBF MAY CONTAIN AN RNA COMPONENT	144

LIST OF FIGURES

PAPER 1

Figure 1. Structure of the 21 kb macronuclear rRNA genes in <i>T. thermophila</i>	16
Figure 2. Footprint analysis of protein-DNA interactions in the Type I repeat	24
Figure 3. Gel mobility shift assays for proteins that bind to Type I repeat DNA	27
Figure 4. Phosphocellulose chromatography of proteins that bind to Type I repeat DNA	29
Figure 5. ds-TIBF interacts with the central A-rich region of the duplex Type I repeat	33
Figure 6. The ssT-TIBF is not sequence specific; the ssA-TIBF recognizes Type I repeats	37
Figure 7. Deletions of the terminal domains of the Type I repeat result in loss of ssA-TIBF recognition	40

PAPER 2

Figure 1. Structure of the <i>Tetrahymena thermophila</i> macronuclear rDNA	71
Figure 2. Purification of ssA-TIBF	79
Figure 3. UV crosslinking of ssA-TIBF with the 37 nt Type I repeat	81

Figure 4. The recognition sequence of ssA-TIBF includes 3' flanking sequences as well as the Type I repeat	85
Figure 5. Additional sequences that enhance ssA-TIBF binding flank two Type I repeats identified as cis-acting replication control elements	87
Figure 6. The binding affinity of ssA-TIBF for C3 vs. B rDNA sequences <i>in vitro</i> parallel their replication efficiencies <i>in vivo</i>	90
Figure 7. ssA-TIBF binds the promoter Type I repeat and upstream repeat sequences with an equal affinity	93

APPENDIX

Figure 1. Nuclease treatments affect the binding of ds-TIBF to the Type I repeat	146
Figure 2. The ds-TIBF-Type I repeat complex is sensitive to RNase H	149

ABSTRACT

Proteins that bind to specific sequences in replication origins have been shown to be important components of both prokaryotic and eukaryotic DNA replication systems. An origin of DNA replication is located within the 5' non-transcribed spacer region of the rRNA genes (rDNA) of *Tetrahymena thermophila*. Previous work has shown that several cis-acting mutations responsible for rDNA replication defects lie in upstream repeats of a transcriptional promoter element, the Type I repeat.

Using FE(II)EDTA cleavage footprinting and gel-mobility shift assays, I identified and characterized two distinct DNA-binding activities that interact with the Type I repeat. One of these activities, designated ds-TIBF, binds preferentially to duplex DNA whereas the other, ssA-TIBF, binds preferentially to single-stranded DNA. Competition studies revealed that although ds-TIBF binds with high affinity to Type I sequences, it has only moderate sequence specificity and binds with high affinity to other duplex oligonucleotides containing dA/T stretches of greater than 6 nucleotides. In contrast, ssA-TIBF does not bind to generally A-rich oligonucleotides, but specifically interacts with single-stranded oligonucleotides corresponding to the Type I repeat.

UV crosslinking of the partially purified protein to A-rich strand oligonucleotides followed by SDS gel electrophoresis demonstrated the binding activity of ssA-TIBF is associated with a polypeptide of

approximately 95 kD. The native size of the binding activity determined by gel filtration chromatography is 205 kD, suggesting that ssA-TIBF may exist as a dimer or that the 95 kD polypeptide associates with another protein(s) in the cell.

The nucleotide sequence requirements for optimal binding of ssA-TIBF were characterized. A deletion that removes part of one copy of the ssA-TIBF recognition sequence in the replication origin region has been shown to cause an rDNA replication defect *in vivo*. The affinity of ssA-TIBF for oligonucleotides reflecting this deletion is reduced by approximately 50% *in vitro*. Another copy of the Type I repeat, located only 20 nucleotides upstream from the transcription start site, is an essential component of the rRNA gene promoter. This element, although slightly different in sequence, is bound by ssA-TIBF with an affinity approximately equal to that of the Type I element in the replication origin region. These findings suggest that ssA-TIBF could be a transcription factor that also plays a role in rDNA replication.

GENERAL INTRODUCTION

One of the ultimate goals in biology is to understand the molecular components necessary to ensure that the genome is completely and accurately reproduced once and only once during the cell cycle.

Although much is known about how prokaryotic organisms replicate their genome, the molecular mechanisms that regulate eukaryotic DNA replication are not yet clear.

Several of the individual components required for eukaryotic DNA replication have been purified and studied using *in vitro* systems. In particular, when this method was used with cellular extracts from Simian Virus 40 infected monkey cell lines, several of the essential cellular factors required for SV40 DNA synthesis were identified (93, 108). The only virally encoded factor required for SV40 DNA replication is the initiation factor large tumor antigen (T antigen)(69, 94, 106). T antigen binds to the SV40 replication origin (96) and unwinds the DNA (92) in preparation for the DNA replication machinery. Individual replication factors purified from whole cell extracts were reconstituted *in vitro* and tested for their ability to replicate plasmids containing the SV40 origin of replication when supplemented with T antigen. Seven cellular factors were determined to be essential for replication using this system (93, 108). These factors include the DNA polymerase α -primase complex, DNA polymerase δ , Proliferating Cell Nuclear Antigen (PCNA), replication factor C (RF-C), replication protein A (RF-A) and DNA topoisomerases.

DNA replication proceeds only in the 5' to 3' direction. DNA synthesis on newly unwound SV40 DNA templates is continuous on one strand, the leading strand, and discontinuous on the other, the lagging strand, creating small RNA primed Okazaki fragments. The DNA polymerase α -primase complex synthesizes an RNA primer and initiates SV40 DNA synthesis from this primer (90). It is thought to be responsible for lagging-strand synthesis, since the primase, tightly associated with the polymerase, is readily available for multiple initiation events. DNA polymerase δ is capable of processive DNA synthesis in the presence of PCNA (82). The processivity of DNA polymerase δ makes this polymerase most likely to be responsible for SV40 leading strand DNA synthesis. RF-C recognizes primed DNA and is required for leading strand synthesis in a reconstituted system (98, 104), therefore RF-C is probably involved in facilitating polymerase δ template binding. RP-A, also known as RF-A or human single-stranded DNA binding protein (hSSB), binds preferentially to single-stranded DNA and facilitates T antigen dependent DNA unwinding (36, 107, 109). E. coli single-stranded binding proteins (SSBs) can substitute for RP-A in template unwinding reactions, but cannot do so in the complete SV40 replication reaction, suggesting RP-A has an additional role in DNA replication probably involving protein-protein interactions. Finally, topoisomerases relieve the torsional stress caused by the unwinding of DNA, and are essential for the dissociation of the replicated daughter circular plasmids (112).

Homologs for DNA polymerases α and δ , PCNA, and RP-A are essential for yeast growth, suggesting a requirement for these proteins for chromosomal DNA replication *in vivo* (reviewed in 16). Another DNA polymerase, DNA polymerase ϵ , is also an essential gene in yeast (76). Polymerase ϵ is highly processive and therefore may also be involved in leading strand synthesis. However, a human homolog to polymerase ϵ is active in DNA repair, suggesting another function for this enzyme (78). Homologs of many of these replication factors have also been identified in other higher eukaryotes (reviewed in 90), suggesting that some of the essential components of DNA replication have been conserved throughout evolution.

Thus, the SV40 *in vitro* system of DNA replication has aided in our understanding of the functional roles of some of the components of the DNA replication machinery once replication has begun. However, since SV40 replication relies on the virally encoded initiation factor T antigen, the actual events initiating eukaryotic DNA replication have not been resolved.

One problem contributing to the difficulty of understanding DNA replication initiation is the lack of genetically-defined DNA sequence elements known to be required for chromosomal replication in eukaryotes. Only in yeast (16, 37), *Drosophila* (17) and *Tetrahymena* (67, 111) have specific sequence elements been implicated in the control of chromosomal DNA replication. In yeast, autonomously replicating sequences (ARS) were originally identified on the basis of their ability to stably maintain plasmids in transformed cells, and have

since been extensively characterized (reviewed in 16, 37). The minimal ARS element consists of an 11 bp AT-rich ARS core consensus sequence (ACS) flanked by an AT-rich region that contains several functional elements, including transcription factor binding sites. More recently, a subset of identified ARS elements have been shown to function in their normal chromosomal locations as replication origins (11, 31, 34, 38, 44, 55, 71, 102, 119). Mutations in ARS core consensus sequences of two chromosomal origins inhibited origin function (31, 83)

DNA-binding proteins that interact with the ARS and may be involved in the initiation of DNA replication have been identified. The ACS and portions of sequences 3' to the T-rich strand of the ACS are protected in both *in vitro* and *in vivo* DNase I footprinting experiments (5, 32). The *in vitro* footprint was created by a purified Origin Recognition Complex (ORC) consisting of 6 polypeptides (5). ORC specifically interacts with the ACS sequence and is sensitive to mutations which reduce the ability of plasmids to replicate autonomously. Another sequence-specific ACS-binding protein (ACSB) has been independently isolated in three different laboratories (54, 65, 86). ACSB specifically recognizes the T-rich single-stranded form of the ACS. The affinity of ACSB for mutant ACS also corresponds to the ability of these sequences to promote plasmid replication. Whether ACSB functions independently of ORC or coordinately is not known.

In contrast to yeast, several other eukaryotes appear to initiate chromosomal DNA replication in broad regions rather than at specific sites. For example, zones of DNA replication initiation found in

mammals (41, 99, 100), *Drosophila* (51, 87), and *S. pombe* (118) span regions from 450 bp to 20 Kbp. It is possible that in these organisms the sites of replication initiation are not determined by specific sequences, but by a combination of several structural features on the chromosomes. Several of these features, such as easily unwound regions, binding sites for transcription factors, and pyrimidine tracts have been found to occur in several examples of DNA replication initiation zones (7, 41).

We have chosen the rRNA genes of the ciliate *Tetrahymena thermophila* as a model system for studying DNA replication because it offers two unique features. First, an origin of DNA replication has been mapped by electron microscopy to the 5' nontranscribed spacer (NTS) region of the rRNA genes (rDNA) (19). Second, *Tetrahymena* rDNA is uniquely amenable to genetic analysis being the only known eukaryote with a single germline copy of the rRNA genes in the micronucleus (reviewed in 113). This has allowed mutants in rDNA amplification and replicative maintenance to be isolated and the cis-acting sequences involved to be identified (67, 111). Therefore, *Tetrahymena* provides a genetically-defined eukaryotic DNA sequence element implicated in the control of DNA replication.

In vegetatively growing *Tetrahymena*, the rDNA is stably maintained as an amplified ($\sim 10^4$ copies), 21kb palindrome in the mature somatic macronucleus (113). The rDNA replicates from an origin 650 $\pm$ 300 bp from the center of the palindromic rDNA molecule within the 5' nontranscribed spacer (5' NTS) (19). A plasmid

containing a 1.9 kb fragment consisting of the 5' NTS is able to stably transform the *Tetrahymena* macronucleus (115). Molecules with multiple copies of the 5' NTS have a replicative advantage over those with fewer copies (116). The origin overlaps with a nuclease hypersensitive region designated Domain 1 and may include an imperfect repeat of this region, Domain 2 (see paper 1, Fig. 1) (18, 80). Also within the 5' NTS are several conserved repeated elements, referred to as Type I, II, and III repeats (20). The Type I repeat located 20 nucleotides upstream of the transcriptional start site is an essential element of the transcriptional promoter *in vitro* (75). The three Type I repeats upstream of the promoter element are not essential for transcription *in vitro*. The role of the Type II repeats, is not yet known, although they form complexes with *Tetrahymena* cellular protein(s) in *in vitro* binding experiments (Ron Pearlman, personal communication). Finally, the Type III repeats are topoisomerase I binding and cleavage sites (10).

Specific sequence elements involved in the replication of the rDNA have been identified by isolating mutants with rDNA replication defects. Natural rDNA variants, C3 and B, that differ in replication ability were utilized. The vegetative progeny of cells crossed to produce heterozygotes containing both C3 and B rDNA, gradually lose their B rDNA (81). Only a 10% difference in replication rates of the two alleles is sufficient to explain the observed loss of B rDNA. C3-*rmm* (replication maturation and maintenance) mutants which reverse the replication preference were isolated (67). Genetic and molecular

analysis revealed that three independently isolated mutations differed from the wild type sequence by a single base pair deletion within two of the upstream Type I repeats (see Paper 1, Fig. 1). Comparison of C3 and B rDNA also showed that C3 and B rDNA differed in a way that affected the Type I repeat. A 42 base pair deletion caused the Type III and Type I repeats in domain 2 to overlap (see Paper 1, Fig. 1). Thus, the Type I repeat is a genetically-defined DNA sequence element implicated in the regulation of rDNA replication.

I have identified several protein factors which interact with the Type I repeat. I show that one of these factors, ds-TIBF, recognizes the AT-rich central region of the duplex repeat. The interactions between ds-TIBF and the Type I repeat may be facilitated by an RNA component. I have also characterized two factors that interact with the single-stranded forms of the Type I repeat. One of these, ssT-TIBF, forms nonspecific interactions with the T-rich single-strand of the Type I repeat. The other, ssA-TIBF, specifically recognizes the A-rich single-strand of the Type I repeat. ssA-TIBF specifically recognizes The Type I repeat and sequences flanking the conserved Type I repeat. These flanking sequences are deleted in B rDNA. The affinity of ssA-TIBF for B and C3 Type I sequences *in vitro* mimic the replication hierarchy that exists *in vivo*, suggesting that ssA-TIBF could mediate the differential replication of the two rDNA alleles *in vivo*. Finally, ssA-TIBF exhibits an equal binding affinity for the Type I repeat that is an essential element of the promoter *in vitro* (75) and for upstream copies of the

Type I repeat. This suggests that ssA-TIBF may be a transcription factor involved in the regulation of rDNA replication.

An explanation of the dissertation organization

The work described in this dissertation has focused on characterizing three DNA-binding factors that interact with a genetically-defined replication control element, the Type I repeat. The dissertation has been divided into two papers and an appendix in accordance with the rules for the alternative thesis format as outlined in the Iowa State University Graduate College Thesis Manual (1992).

The first paper describes the initial characterization of the specificity with which several factors interact with the Type I repeat. This multiple author paper has been submitted for publication to *Molecular and Cellular Biology*. The entirety of the first paper was done by the doctoral candidate, except for the experiment depicted in Figure 2. The paper was written jointly by the doctoral candidate and the principal investigator, Drena Larson.

The second paper describes the purification of one of these factors, ssA-TIBF, further characterizes the sequences bound by ssA-TIBF, and defines the potential role of ssA-TIBF in rDNA replication. This paper will also be submitted to *Molecular and Cellular Biology*. The research in the second paper was completed entirely by the doctoral candidate. The paper was written by the doctoral candidate with editorial contributions from Drena Larson and Robert Benbow.

The appendix describes preliminary work suggesting that another of the factors, ds-TIBF, may contain an RNA component. The research within the appendix was done entirely by the doctoral candidate. The section was written by the doctoral candidate with editorial contributions from Drena Larson.

Preceding the papers I have written a general introduction. Following the papers are a general summary of conclusions, including suggests for future direction, and the references cited in the general introduction, general conclusions and appendix. The appendix follows the list of references cited.

PAPER 1. A SINGLE-STRANDED DNA-BINDING PROTEIN BINDS
SPECIFICALLY TO A REPLICATION CONTROL ELEMENT IN THE
ORIGIN REGION OF *TETRAHYMENA THERMOPHILA* rDNA

**A Single-Stranded DNA-Binding Protein Binds Specifically to a
Replication Control Element
in the Origin Region of *Tetrahymena thermophila* rDNA**

Angela R. Umthun, Wen-Ling Shaiu, Zita A. Sibenaller[†], Drena D. Larson*

Department of Zoology and Genetics
Signal Transduction Training Group
Molecular, Cellular and Developmental Biology Program
Interdepartmental Genetics Program

Room 2114 Molecular Biology Building
Iowa State University, Ames, IA 50011

* Corresponding author.

Phone: (515)294-1112

FAX: (515)294-2876

Email: ddlaron@molebio.iastate.edu

[†] Present address; Molecular Biology Program, College of Medicine,
University of Iowa, Iowa City, IA 52242

ABSTRACT

Mutations in AT-rich conserved Type I repeat sequences located in the replication origin region of the rRNA genes (rDNA) in *Tetrahymena thermophila* affect the ability of rDNA molecules to replicate in macronuclei (35). Fe(II) EDTA cleavage footprinting of restriction fragments containing a Type I repeat showed that most of the 33 conserved nucleotides were protected by proteins in a S100 supernatant of a *Tetrahymena* whole cell extract. Gel mobility shift assays were used to characterize DNA-binding proteins interacting with the Type I repeat sequence. Multiple proteins that bound to the Type I repeat were identified using a double-stranded synthetic 33-bp oligonucleotide substrate. One of these, designated ds-TIBF, bound preferentially to duplex DNA with high affinity, but with only moderate specificity for Type I sequences. Using the T-rich strand of the Type I repeat as a substrate, a protein(s) that bound nonspecifically to T-rich single-stranded DNA was identified. In contrast, a single-stranded DNA-binding protein(s) that bound the A-rich strand of the Type I repeat, designated ssA-TIBF, specifically recognized single-stranded oligonucleotides containing the Type I repeat sequence. Deletion of the 5' or 3' borders of the conserved sequence greatly reduced the binding of the ssA-TIBF. The binding properties of ssA-TIBF *in vitro*, coupled with the observation that mutations in the Type I sequence affect the ability of mutant rDNA molecules to replicate *in vivo*, suggest a possible physiological role for ssA-TIBF in rDNA replication in *Tetrahymena*.

INTRODUCTION

Chromosomal origins of DNA replication in eukaryotes have been extraordinarily difficult to isolate and define (reviewed in reference 4). Despite tremendous effort, however, specific cis-acting DNA sequences required for origin activation been identified only in the yeast, *S. cerevisiae* (reviewed in references 9, 18). Yeast autonomously replicating sequences (ARSs) were originally isolated on the basis of their ability to promote extrachromosomal replication of plasmids in yeast. Mutational analyses have shown that the DNA sequences essential for ARS activity on plasmids include an 11 bp ARS core consensus sequence (ACS) and an AT-rich flanking region that contains several functional elements, including binding sites for transcription factors (39, 64) and DNA unwinding elements (DUEs)(60). In two different chromosomal origins, specific mutations within the ARS core consensus sequence were shown to abolish origin function (15, 45). The ACS element in yeast is the only genetically-defined DNA sequence element known to be required for chromosomal DNA replication in eukaryotes.

Several DNA binding proteins that recognize the ACS or other functional elements in ARS sequences have been identified. Three laboratories characterized a single-stranded DNA binding protein that binds preferentially to the T-rich strand of the ACS (29, 33, 34, 49). In addition, a multiprotein complex that interacts with the ACS in double-stranded DNA was purified (3). This origin recognition complex generates nuclease protection patterns over the ACS *in vitro* that are

very similar to those observed *in vivo* in high resolution genomic footprinting experiments (16).

Specific cis-acting sequences that regulate rDNA replication in *Tetrahymena thermophila* were identified (35, 69). The rDNA in *Tetrahymena* is present as a single chromosomal copy in the germline micronucleus and as highly amplified 21kb linear palindromic replicons in the somatic macronucleus. rDNA amplification (to $\sim 10^4$ copies) occurs during macronuclear development in mating cells; after amplification, the rDNA is stably maintained at this high copy number during vegetative division of cells (reviewed in reference 70). The amplified rDNA replicates from a bidirectional origin within the 5' nontranscribed spacer region (5'NTS) upstream from the rRNA transcription initiation site (Fig. 1) (11). A 1.9 kb DNA fragment encompassing the 5'NTS promotes autonomous replication of circular plasmids and short linear chromosomes in the *Tetrahymena* macronucleus (71). Molecules containing multiple copies of the origin region had a replication advantage over those with fewer copies (72).

The 5'NTS contains three families of evolutionarily conserved repeated sequences referred to as Type I, Type II and Type III repeats (12, 37) and three nuclease hypersensitive regions (42) (Fig. 1). Genetic and molecular studies of *Tetrahymena* rDNA alleles with differential replication properties identified the Type I repeat as a cis-acting replication control element (43). The rDNA allele of inbred strain C3 had a replication advantage over that of inbred strain B when both alleles reside in the same macronucleus. Several *rmm* (rDNA maturation

Fig. 1. Structure of the 21 kb macronuclear rRNA genes in *T. thermophila*.

(A) A schematic diagram of the 21 kb palindromic rDNA. Arrows indicate rRNA transcription units and hatched lines represent telomeric sequences. The 5' and 3' non-transcribed spacer regions (5'NTS and 3'NTS) are indicated.

(B) An enlarged diagram of the 5'NTS. Broad vertical arrows indicate DNAase hypersensitive domains (42). Putative replication origin(s) are shown (10, 11, 42). Domains 1 and 2 refer to DNase hypersensitive domains that correspond to an ~400 bp imperfect duplication.

(C) A schematic diagram of repeated sequence elements in the 5'NTS of C3 rDNA. Conserved sequence elements designated Type I, II and III repeats are indicated (12). Asterisks denote Type I elements known to be altered in mutants defective in rDNA replication (35, 51, 69).

21 kb rDNA

and maintenance) mutations that abolish the replication advantage of C3 rDNA were isolated and characterized (35, 51, 69). The replication defects in strain B and in three independently isolated C3-*rmm* mutants result from sequence changes that alter copies of the Type I repeat within the 5'NTS. It was proposed that competition for limiting trans-acting factors that bind to the Type I repeats could explain the differential replication properties of rDNA alleles, and that Type I repeats could be involved in the initiation of rDNA synthesis (35).

We report here the characterization and partial purification of three distinct DNA-binding proteins from *Tetrahymena* that recognize the Type I repeat. One of these binds with high affinity but only moderate specificity to the duplex form of the repeat. The others bind to single-stranded DNA, one non-specifically to T-rich DNA and the other specifically to the A-rich strand of the Type I sequence. The interaction of these proteins with the Type I repeat, an rDNA replication control sequence, suggests that one or all may play a role in the regulation of rDNA replication *in vivo*.

MATERIALS AND METHODS

Cell growth and S100 extract preparation

S100 extracts from *Tetrahymena* were prepared essentially as described by Greider and Blackburn (26). *T. thermophila* strain C3V cells were grown at 30°C in 2% PPYS (2% proteose peptone, 0.2% yeast extract, 0.003% sequesterine) to a density of 2.5×10^5 cells/ml. Cells were collected by centrifugation at $2600 \times g$ for 7 min, washed twice in 10mM Tris-HCl, pH 7.5, and resuspended in HMG (20 mM HEPES, pH 7.9, 1 mM $MgCl_2$, 10% v/v glycerol) in 5X the pellet volume. Cells were lysed by the addition of 0.2% NP40 in HMG for 30 min on ice, while stirring. The whole cell extract was centrifuged at $100,000 \times g$ for 60 min in a Beckman TLA100.3 fixed angle rotor. Aliquots of the supernatant (S100) were frozen immediately in liquid nitrogen, and stored at -80°C.

Hydroxyl radical footprinting

Substrates for hydroxyl radical footprinting were derived from plasmid pUC3-XN, which contains a 770 bp *Xba*I to *Nsi*I restriction fragment from the rDNA of strain C3V cloned into pUC118. This rDNA fragment contains the region designated Domain 2 in Fig. 1. For analysis of the A-rich strand of the Type I repeat, *Xba*I digested pUC3-XN DNA was 5' end-labeled using T4 polynucleotide kinase and [γ - ^{32}P]dATP, and subsequently digested with *Dra*I. For analysis of the T-rich strand of the Type I repeat, pUC3-XN DNA was linearized at the *Kpn*I site in the polylinker (adjacent to the *Xba*I site), 3' end-labeled using terminal

deoxytransferase and [^{32}P]ddATP, and subsequently digested with *Dra*I. Labeling and gel purification of DNA fragments were carried out according to standard procedures (47).

The Fe(II) EDTA cleavage reactions were performed according to the method of Tullius and Dombroski (59). Binding reactions (70 μl) were performed as described for gel mobility shift assays (see below) and incubated 10 min on ice. 10 μl Fe(II)EDTA (equal volumes of 0.2mM Fe(II){ $(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ } and 0.4 mM EDTA mixed just before use) was added to the binding reaction with 10 μl 0.3% H_2O_2 and 10 μl 10mM sodium ascorbate. The reaction was incubated at room temperature for 5 min and stopped by addition of 10 μl of 0.1 M thiourea, 32 μl of 0.2 M EDTA, and 10 μg of tRNA. After ethanol precipitation, the cleavage products were separated on a 10% sequencing gel.

Gel mobility shift assays

Most substrates for gel mobility shift assays were oligonucleotides synthesized by the Nucleic Acid Facility, Iowa State University. The *T. thermophila* Type II oligonucleotides were a gift from R. Pearlman, York University. Single-stranded oligonucleotides were 5' end-labeled using T4 polynucleotide kinase and purified using Select D spin columns (5' Prime-3' Prime, Boulder, CO). To form duplex oligonucleotides, equal molar amounts of labeled complementary strands were mixed in 100 mM NaCl, heated to 70°C for 10 min and allowed to cool to room temperature.

Gel mobility shift assays were performed essentially as described (20, 24). In standard reactions, 1 ng of ^{32}P -labeled

oligonucleotide and 1 μ g poly(dI-dC) were incubated in the presence of 5 μ l S100 extract (17.5 μ g protein) or partially purified fractions for 15 min on ice. The final reaction mixture contained 17 mM HEPES, pH 7.9, 1 mM EDTA, 200 mM NaCl, 8.7% glycerol, 1 mM DTT, 0.087 mM PMSF in a total of 15 μ l. For competition assays, 1 μ l of the appropriate concentration of unlabeled oligonucleotide in STE (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA), was added to the sample prior to the addition of extract. Samples were subjected to electrophoresis in 5% or 10% polyacrylamide gels in 0.6X Tris-borate-EDTA (47) at 200V (<50amps) for 60 min. After electrophoresis, gels were dried and exposed to Kodak XAR-5 film or phosphorImager screens. Radioactivity in individual bands was quantitated using a Molecular Dynamics PhosphorImager.

Column chromatography

For separation and partial purification and of Type I binding factors, 0.4 ml of *T. thermophila* S100 extract (4.8 μ g/ml protein) was applied to a phosphocellulose column (P11, Whatman, 1 x 4 cm) equilibrated with TEG/0.1 (25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 10% glycerol, 0.02% NP40, 1 mM DTT, 0.1 mM PMSF, 0.1 M NaCl). The column was washed with 15 ml TEG/0.1 and bound proteins eluted with a step gradient of NaCl (0.2M, 0.3M, 0.4M, 0.6M, and 1M) in TEG. The elution position of Type I binding proteins was determined using gel mobility shift assays. Protein concentrations were determined by the method of Bradford (6).

Gel filtration chromatography was used as an alternative procedure to separate Type I binding factors and to estimate their apparent native molecular weights. 1 ml of S100 extract (3.5 μ g/ml) was applied to a Sephacryl S-300 column (Pharmacia, 1.5 x 88 cm) equilibrated with TEG/0.2 (TEG with 0.2 M NaCl). Proteins were eluted with 200 ml of the same buffer and 1 ml fractions were collected. Gel mobility shift assays were used to determine the elution positions of Type I binding proteins. Stokes radii of the binding proteins and estimated native molecular weights (assuming globular proteins) were determined by comparison with molecular weight calibration standards (Boehringer-Mannheim Biochemicals, Inc., Indianapolis, IN). Sizes reported were the average of values obtained in two experiments.

RESULTS

Identification of DNA binding proteins that recognize Type I repeat sequences.

Hydroxyl radical footprinting was used to characterize DNA-protein interactions associated with Type I repeat sequences in the replication origin region of *T. thermophila* rDNA. Restriction fragments containing the Type I element from the nuclease hypersensitive region designated Domain 2 (Fig. 1) were used as substrates for Fe(II)EDTA cleavage reactions. These experiments revealed a complex pattern of protein-DNA complexes in the vicinity of the Type I element (Fig. 2). Most nucleotides within the 33 bp conserved Type I element were protected from cleavage, including 25 nt on the A-rich strand (upper strand in Fig. 2B) and 22 nt on the T-rich strand. Three mutations that cause rDNA replication defects *in vivo* result from the deletion of one base pair in the eleven central A-T base pairs of the Type I repeat (35, 55, 69). Seven of these positions were protected from cleavage on both strands, whereas the remaining four positions were protected only in the T-rich strand. Other protected regions in sequences flanking the Type I element on both sides were observed, as were several positions of apparently enhanced cleavage both within the Type I sequence and in flanking regions.

Gel mobility shift assays were used to investigate the sequence specificity of the Type I repeat binding factor(s) (TIBFs) detected in footprinting experiments. These assays revealed one major shifted band when the same rDNA restriction fragment used for footprinting was

Fig. 2. Footprint analysis of protein-DNA interactions in the Type I repeat.

(A) Fe(II)EDTA cleavage footprinting of restriction fragments containing a Type I repeat.

Restriction fragments containing a single Type I repeat from C3 rDNA were ³²P-labeled on either the A-rich (5'-end-labeled) or T-rich (3'-end-labeled) strand and incubated with Fe(II)EDTA in the presence (+) or absence (-) of *Tetrahymena* whole cell extracts as described in Materials and Methods. Maxam-Gilbert chemical sequencing reactions for each fragment are shown in G+ A lanes. Open boxes on the right show the position of the Type I repeat within the fragment. Brackets on the left indicate the portion of the DNA sequence shown in Fig. 2B.

(B) Schematic representation of DNA sequences protected from Fe(II)EDTA cleavage in *Tetrahymena* whole cell extracts. The DNA sequence of the Type I repeat (boxed) and surrounding sequences in Domain 2 of C3 rDNA are shown. The top strand is the A-rich strand; the bottom strand is the T-rich strand. Cross-hatched boxes denote protected nucleotides on each strand.

B

```

      |||  |||  |||  |||  |||  |||
5' CATTAGAAA TTTTTTGGCAAAAAAAAAA CAAAAATAGTAA ACCTTCCGAA 3'
3' GTAATCTT TAAAAAACCGTTTTTTTTTGT TTTTATCAT TGAAGGCTT 5'
      |||  |||  |||  |||  |||  |||

```

employed as the binding substrate (data not shown). Since synthetic oligonucleotides corresponding to the Type I repeat competed strongly for binding to this rDNA segment, Type I repeat-containing oligonucleotides were used as substrates in subsequent experiments presented here. A 33 bp double-stranded oligonucleotide containing only the Type I repeat sequence (the boxed sequence in Fig. 2B) bound a factor(s) in an S100 extract to produce a single major shifted band (Fig. 3). In this experiment, an 8-fold molar excess of unlabeled Type I oligonucleotide prevented formation of the major shifted complex, whereas a 40-fold excess of an unrelated sequence did not. These results suggest that the Type I repeat is recognized specifically by DNA-binding factors in *Tetrahymena* cellular extracts.

Electrophoresis on 5% polyacrylamide gels resolved the single Type I DNA-protein complex into multiple species (data not shown). Since duplex Type I oligonucleotides probes usually contain a low level of unhybridized single-stranded oligonucleotides, we investigated whether the multiple species could result from single-stranded as well as double-stranded DNA-binding proteins.

Separation of duplex and single-stranded Type I binding factors

Single-stranded and double-stranded DNA-binding proteins were resolved and partially purified by phosphocellulose column chromatography (Fig. 4). The binding proteins were eluted with a mixture of a double-stranded Type I oligonucleotide and the

Fig. 3. Gel mobility shift assays for proteins that bind to Type I repeat DNA.

Complexes between factors in S100 supernatants of whole cell extracts and a Type I repeat oligonucleotide were formed in the absence (0) or presence of increasing amounts of unlabeled competitor DNA and visualized on a 10% polyacrylamide gel. Binding reactions contained 1 ng of ^{32}P -labeled duplex 33 bp Type I repeat sequence, the indicated molar amounts of competitors, 1 μg poly dI-dC, and 3 ml S100 extract (see Materials and Methods). Lanes 2-6, 0, 1X, 2X, 4X or 8X molar excess of duplex Type I competitor ; lanes 7-11, 0, 5X, 10X, 20X, or 40X of the nonspecific duplex competitor (5'-AACCCGATCACCCC-3'). Input indicates labeled Type I repeat duplex DNA with no extract added; B, bound DNA (protein-DNA complexes); U, unbound input DNA.

Fig. 4. Phosphocellulose chromatography of proteins that bind to Type I repeat DNA.

S100 extracts were fractionated by phosphocellulose chromatography as described in Materials and Methods. The indicated fractions were assayed for ability to form specific nucleoprotein complexes with a mixed probe containing duplex Type I repeat and both the A-rich and T-rich single-strands of the Type I repeat. Load indicates S100 extract prior to fractionation; ss and ds denote complexes subsequently determined to contain proteins that bind preferentially to single-stranded or duplex DNA. The two large complexes present only in the load and wash lanes probably resulted from aggregation.

unhybridized single-stranded A-rich and T-rich oligonucleotides. Four major DNA-protein complexes were identified (Fig. 4). By reassaying the column fractions with each oligonucleotide separately, three of these were identified as single-stranded Type I DNA binding factors (ss-TIBFs) and one was determined to be a double-stranded Type I DNA binding Factor (ds-TIBF).

The ds-TIBF eluted between 0.3 and 0.4 M NaCl (Fractions 31-39). The higher molecular weight ss-TIBFs were a mixture of proteins that bound the A-rich oligonucleotide (ssA-TIBFs), and the T-rich oligonucleotide (ssT-TIBF). The lower molecular weight ss-TIBF bound preferentially to the T-rich oligonucleotide (ssT-TIBF). The ssTIBFs were subjected to gel-filtration chromatography on Sephacryl S300, and the native molecular weights of the binding proteins were estimated. The Stokes radius of the ds-TIBF was 38Å, corresponding to an apparent molecular weight of 105 kD. The three ssA-TIBFs were calculated from the Stokes radii to correspond to molecular weights of 250kD, 190kD, and 180kD, and the four T-TIBFs were calculated to correspond to molecular weights of 250kD, 200kD, 180kD, and 50kD. These results were in agreement with the apparent sizes of the protein complexes in gel mobility shift assays (e.g., Fig.4).

Interaction of the duplex Type I binding factor (ds-TIBF) with the AT-rich central region of the Type I repeat

To further characterize the sequence specificity of ds-TIBF, we carried out competition experiments using ~200-fold purified ds-TIBF from the pooled phosphocellulose column fractions. In these

experiments, the ^{32}P -labeled input DNA was the 33 bp double-stranded Type I repeat oligonucleotide. Unlabeled competitor DNAs corresponding to altered versions of the Type I repeat or unrelated nonspecific sequences are shown in Fig. 5A. These included altered Type I repeat sequences in which either 1 or 5 A-T base pairs in the central A_{11} tract were deleted (A_1 and A_5 deletions) or two C-G base pairs were substituted for two G-C base pairs, unrelated Type II repeat sequences derived from the 5'NTS region of the rDNA from either *T. thermophila* (see Fig. 1) or *T. pyriformis*, and a 33 bp homopolymer, oligo(dA-dT)₃₃. Fig. 5B is a representative gel mobility shift assay in which binding of ds-TIBF to the Type I repeat substrate was challenged with increasing amounts of unlabeled competitor DNAs.

Alterations of the Type I repeat had little effect on the ability of ds-TIBF to bind to the input DNA under standard gel shift assay conditions (Fig. 5C). The A_1 deletion, which mimics the deletion found in three rDNA replication mutants (35, 55, 69) or deletion of 5As had no statistically significant effect. A competitor containing substituted nucleotides bordering the central A-rich region also bound ds-TIBF (Fig. 5C). Some feature of the duplex Type I sequence appeared to be important for binding, however, since several random sequence oligonucleotides did not compete (Fig. 3 and data not shown). The apparent specific interaction of ds-TIBF with Type I repeat sequences appeared to be due to an affinity for long tracts of A-T base pairs, since competition by oligo(dA-dT)₃₃ was equivalent to that of the Type I repeat oligonucleotide (Fig. 5B and C). Nonspecific competitor DNAs

Fig. 5. ds-TIBF interacts with the central A-rich region of the duplex Type I repeat.

Gel mobility shift analysis of Type I repeat binding by phosphocellulose-purified ds-TIBF in the presence of oligonucleotide competitors was carried out as described in Materials and Methods.

(A) Synthetic oligonucleotides used as competitors against Type I repeat oligonucleotides in gel mobility shift assays. Shaded regions correspond to the central A₁₁ tract from which 1 or 5 As were deleted. Lower case letters indicate base substitutions. The ability of each oligonucleotide to compete for binding to ds-TIBF relative to the unaltered duplex Type I repeat at a 50X molar excess is indicated; + denotes competition level $\geq 75\%$ and - denotes competition level $\leq 25\%$ of the value obtained with the unaltered duplex Type I repeat.

(B) Representative gel mobility shift experiment showing that ds-TIBF binding is resistant to competition by single-stranded Type I sequences but sensitive to competitors containing long runs of duplex AT-rich sequences. Concentrations of unlabeled competitors were: 0, 10X, and 50X molar excess for the duplex Type I repeat (Type I_{DS}) and duplex (dA-dT)₃₃ competitors; 0, 15X, and 75X for the duplex Type II oligonucleotide; 0, 20X, and 100X for the A-rich (Type I_A) and T-rich (Type I_T) single-stranded competitors.

(C) Summary of gel mobility shift experiments using duplex competitors. Sequences of the indicated competitors are shown in Fig. 5A. Radioactivity in Type I DNA-protein complexes was quantitated using a Molecular Dynamics PhosphorImager. The value for the Type I repeat complex in the absence of competitor was taken as 100% bound, and other values are expressed as a percentage of this value. Each data point represents an average of 2 to 3 independent experiments, except that for the G to C substitution at 50X, which is derived from a single experiment. These results were corroborated by 7-9 independent experiments in S100 extracts (data not shown). Standard deviations are indicated for each data point unless the deviation bars fall within the point symbol.

(D) Summary of gel mobility shift competition experiments comparing duplex and single-stranded oligonucleotides. Values were determined as described for Fig. 5C.

A Synthetic Oligonucleotides:		Competition ability
Duplex Type I repeat		
Type I repeat	TTTTTTTGGCAAAAAAAAAACAAAATAGTAA AAAAAAACCGTTTTTTTTTGTTTTTATCATT	+
A ₁ deletion	TTTTTTTGGCAAAAAAAAAACAAAATAGTAA AAAAAAACCGTTTTTTTTTGTTTTTATCATT	+
A ₅ deletion	TTTTTTTGGCAAAAAAAAAACAAAATAGTAA AAAAAAACCGTTTTTTTTTGTTTTTATCATT	+
GG to CC substitution	TTTTTTTccCAAAAAAAAAACAAAATAGTAA AAAAAAaggGTTTTTTTTTGTTTTTATCATT	+
Nonspecific Competitors		
Type II:	GATCCTTAAAAATGAGTGGG	-
<i>T. thermophila</i>	GAATTTTACTCACCTCTAG	
Type II:	TCCACTCAAAAAAGTGAGTGAG	-
<i>T. pyriformis</i>	AGGTGAGTTTTTCACTCACTC	
ds (dAdT) ₃₃	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	+

corresponding to the Type II repeat sequence (see Fig. 1) did not compete for binding (Fig. 5B and C). These oligonucleotides contain a central AT-tract of five or six base pairs (Fig. 5A). Their inability to compete with the Type I oligonucleotide for binding is not simply the consequence of their shorter length, since (dA-dT)₂₅ competes effectively for binding (data not shown). These results suggest that ds-TIBF interacts with tracts of A-T base pairs greater than six base pairs in length and recognizes the A-rich central region of the Type I repeat. Small insertions in the AT-tracts appear to be tolerated, however, because the oligonucleotide in which 5 out of 11As in center of the Type I repeat were deleted still bound ds-TIBF even though the remaining A₁₁-tract was interrupted by a single G-C base pair (Fig. 5A, and B).

AT-rich DNA binding proteins that recognize a 17 bp AT-tract in the SV40 origin of replication have been shown to bind to both double-stranded and single-stranded DNA (23, 38, 57). Therefore, we examined the binding of ds-TIBF to single-stranded oligonucleotides in quantitative competition experiments. Each duplex oligonucleotide tested bound ds-TIBF with at least 5-fold greater affinity than the corresponding the single-stranded competitor (Fig. 5D). Together with the results described above, these experiments demonstrate that ds-TIBF exhibits a strong preference for AT-rich double-stranded DNA, but has only moderate specificity for the Type I repeat sequence.

Interaction of the single-stranded Type I binding factors ssT-TIBF and ssA-TIBF with T-rich and A-rich oligonucleotides

The sequence-specificity of the ssT-TIBFs was first evaluated using S100 extracts. None of the DNA binding proteins detected by the T-rich single-stranded input DNA appeared to recognize the Type I repeat in a sequence-specific manner, since every T-rich oligonucleotide tested competed for binding (Fig. 6A and data not shown). In contrast, the proteins that bound A-rich single-stranded probe specifically recognized the Type I repeat sequence. As shown in Fig. 6B, the A-rich Type I oligonucleotide competed strongly for ssA-TIBF binding, whereas neither (dA)₃₃ nor a nonspecific oligonucleotide corresponding to the A-rich strand of the Type II repeat were effective competitors. Although the relative amounts of the three protein-DNA complexes detected using the A-rich Type I sequence input DNA varied depending on the extract or column fraction being assayed, all three bands appeared and disappeared coordinately in every competition experiment, as well as in extract titration experiments (data not shown).

Sequence-specific interactions of ssA-TIBF with the Type I repeat

To further investigate the sequence specificity of ssA-TIBF, we performed quantitative binding competition assays using pooled Sephacryl S-300 fractions enriched in the largest ssA-TIBF activity (~250kD). In these experiments, the ³²P-labeled input DNA was a 33 nt single-stranded oligonucleotide corresponding to the A-rich strand of

Fig. 6. The ssT-TIBF is not sequence specific; the ssA-TIBF recognizes Type I repeats.

Gel mobility shift assays with 5 μ l of S100 extract binding to either the T-rich (panel A) or A-rich (panel B) single strand of the Type I repeat. The triangles signify increasing concentrations of unlabeled specific and nonspecific single-stranded competitors.

(A) Labeled T-rich Type I repeat (ssT Type I) in competition with 0, 10X and 50X molar excess of the indicated competitors: ssT Type II, T-rich strand of the Type II repeat; ss dT₃₃, oligo(dT)₃₃; ssT Type I, T-rich strand of the Type I repeat.

(B) Labeled A-rich Type I repeat (ssA Type I) in competition with 0, 10X and 50X molar excess of the indicated competitors: ssA Type II, A-rich strand of the Type II repeat; ss dA₃₃, oligo(dA)₃₃; ssA Type I, A-rich strand of the Type I repeat.

the Type I repeat. Single-stranded competitor oligonucleotides corresponding to altered versions of the A-rich strand are shown in Fig. 7A. These include a GG to CC substitution, an A₁ deletion and Type I sequences truncated on either the 5' or 3' end. A representative gel mobility shift assay testing ability of oligonucleotides to compete with the A-rich single strand of the Type I repeat for binding to ssA-TIBF is shown in Fig. 7B. Deletion of a single A or 5A nucleotides (not shown) from the A₁₁ tract did not significantly impair binding to ssA-TIBF (Fig. 7C). However, removal of all sequences flanking the central A-rich region on the 5' side resulted in a significant reduction in binding, as did deletion of the six nucleotides that form the 3' end of the repeat (Fig. 7B, and C). Oligonucleotides bearing these deletions competed only as well as the nonspecific (dA)₃₃ competitor. Curiously, neither the deletion of seven Ts on the 5' end nor the substitution of two Cs for Gs between the T₇ and A₁₁ tracts had any effect on ssA-TIBF binding, even though the oligonucleotide lacking all of these nucleotides (T₇GGC deletion) competed poorly for binding. ssA-TIBF binding ability was not determined by the length of the oligonucleotides used in these experiments, because a 26 nt oligonucleotide (T₇ deletion) competed effectively whereas a 27 nt oligonucleotide (TAGTAA deletion) did not. Taken together, these results demonstrate that ssA-TIBF is a single-stranded DNA binding protein that specifically recognizes the Type I repeat sequence.

Fig. 7. Deletions of the terminal domains of the Type I repeat result in loss of ssA-TIBF recognition.

Gel mobility shift assays of Type I repeat binding by partially purified ssA-TIBF (Sephacryl S-300 fraction) in the presence of oligonucleotide competitors was carried out as described in Materials and Methods.

(A) Synthetic oligonucleotides used as competitors against the A-rich single-strand of the Type I repeat in gel mobility shift assays. Boxed sequences correspond to Type I repeat sequences; shaded regions indicate the central A₁₁ tract; lower case letters denote base substitutions. The ability of each oligonucleotide to compete for binding to ss-TIBF relative to the unaltered A-rich strand of the Type I repeat at a 50X molar excess is indicated.

(B) Representative gel mobility shift experiment showing that ssA-TIBF binding is sensitive to sequence alterations in the terminal portions of the Type I repeat. Concentrations of unlabeled competitors: 0, 10X, and 50X molar excess for each oligonucleotide except the T₇GGC deletion in which 0, 15X and 75X concentrations were used. Sequences of the oligonucleotides indicated above each triangle are shown in Fig. 7A.

(C) Summary of gel mobility shift competition experiments using single-stranded oligonucleotides. Sequences of the indicated competitors are shown in Fig. 7A. Radioactivity was quantitated and values determined as described in the legend for Fig. 5C. Each point represents an average of 6 to 12 independent experiments except for the following competitors: A₁ deletion and TAGTAA deletion, 2 experiments each; T₇GGC deletion, 4 experiments. Standard deviations are indicated for each point unless the deviation bars fall within the point symbol.

A Synthetic Oligonucleotides:		Competition ability (%)
Single-stranded A-rich Type I repeat		
Type I repeat	TTTTTTGGCAAAAAAAAAACAAAAATAGTAA	100
GG to CC substitution	TTTTTTccCAAAAAAAAAACAAAAATAGTAA	107
A ₁ deletion	TTTTTTGGCAAAAAAAAAACAAAAATAGTAA	102
T ₇ deletion	GGCAAAAAAAAAACAAAAATAGTAA	92
T ₇ GGC deletion	AAAAAAAAACAAAAATAGTAA	69
TAGTAA deletion	TTTTTTGGCAAAAAAAAAACAAAA	56
Nonspecific Competitor		
ss (dA) ₃₃	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	53

DISCUSSION

The only genetically-defined DNA sequence elements directly implicated in the control of chromosomal DNA replication in eukaryotes are ARS sequences in *S. cerevisiae* (reviewed in 9, 18), amplification control elements (ACE) and amplification enhancing regions (AER) in the chorion genes of *D. melanogaster* (reviewed in 41), and Type I repeat sequences in the macronuclear rDNA of *T. thermophila* (35, 69). The ARS core consensus sequence is essential for the activation of both chromosomal and plasmid replication origins in yeast (15, 45). The core sequence has long been thought to play a role in the initiation of DNA replication, and DNA-binding proteins that recognize it have been recently identified (3, 29, 33, 34, 49). The ACE and AER are cis-acting control elements necessary for amplification of chorion gene clusters on the third and X chromosomes in *D. melanogaster*. Sites at which DNA replication initiates during amplification have been mapped to regions containing these elements by two-dimensional gel analysis (13, 27).

Much less is known about the role of Type I repeat sequences in *Tetrahymena* rDNA replication. The origin of replication of the palindromic rDNA has been mapped by electron microscopy to a region 650 ± 300 bp from the center of the palindrome (11)(Fig. 1). The mapped origin coincides with a nuclease hypersensitive domain designated Domain 1. The DNA sequence and nuclease sensitivity of this region are duplicated in Domain 2, which may also function as an origin (42). The presence of evolutionarily conserved Type I repeats in the origin region and the fact that mutations in Type I repeats affect the

ability of rDNA molecules to replicate *in vivo*, suggest the hypothesis that Type I repeats are involved in the initiation of rDNA synthesis. To investigate the possible role of the Type I repeat in rDNA replication, we categorized three families of DNA-binding proteins that recognize the sequence *in vitro*: ds-TIBF (Fig. 5), ssT-TIBF (Fig. 6), and ssA-TIBF (Fig. 6 and Fig. 7). One of these, ssA-TIBF, specifically interacts with the A-rich single strand of the Type I repeat and binds with reduced affinity to truncated versions of the conserved sequence element. These results suggest that DNA-protein complexes formed between ssA-TIBF and the Type I repeats in the origin region may play a role in *Tetrahymena* rDNA replication *in vivo*.

ds-TIBF

Several families of DNA-binding proteins that interact with Type I repeat sequences were identified by hydroxyl radical footprinting and gel mobility shift analyses. The double-stranded DNA-binding protein, ds-TIBF, binds with high affinity but only moderate specificity to the duplex form of the repeat. The affinity of ds-TIBF for oligo(dA-dT)₃₃ implies that AT-richness is an important determinant for binding. Type II repeat sequences from the 5'NTS of the rDNA (Fig. 1C), however, are not bound by ds-TIBF even though they have a central AT-tract of 5-6 base pairs (Fig. 5A). One interpretation of these results is that ds-TIBF binding requires tracts of more than 6 A-T base pairs for binding. Alternatively, ds-TIBF could specifically recognize a structural feature of the 17 bp AT-rich central region of the Type I repeat that is also present in oligo(dA-dT)₃₃.

AT-rich duplex DNA binding proteins exist in other eukaryotes. For example, a protein, a 10 kD high mobility group (HMG) protein from monkey cells, binds preferentially to any tract of at least 6 AT base pairs in duplex DNA (53). Several mammalian cellular proteins specifically recognize the 17 bp AT-tract within the SV40 core origin of replication (2, 23, 38, 57). In contrast to ds-TIBF, which exhibits a strong preference for duplex DNA, most of these proteins bind to both single-stranded and duplex DNA.

The DNA-binding activity of ds-TIBF is associated with a 66kD polypeptide (data not shown). It is, therefore, distinct from the much smaller HMG B and HMG C proteins characterized in *Tetrahymena* (50). Although the binding of ds-TIBF to the Type I repeat *in vitro* is only moderately sequence-specific, our experiments do not exclude the possibility that ds-TIBF could selectively interact with the Type I repeat *in vivo*, perhaps by its association with other cellular proteins. It is also possible that ds-TIBF specifically recognizes a conformational feature of the Type I repeat sequence in native chromatin.

ssT-TIBF

Single-stranded DNA-binding proteins detected using as substrate the T-rich strand of the Type I repeat bound nonspecifically to any T-rich single-stranded DNA tested. Three abundant single-stranded DNA binding proteins have been previously identified in *Tetrahymena* (58), but they are considerably smaller than the T-rich binding proteins identified in this study. Single-stranded DNA binding proteins with

relatively little sequence specificity have been shown to play a role in DNA replication in prokaryotes as well as in SV40 DNA replication in mammalian cell extracts. Replication protein A (RP-A), also known as replication factor A (RF-A), or single-stranded DNA-binding protein (SSB) in humans, is essential for SV40 DNA replication *in vitro* (17, 67, 68). The yeast and human RP-A proteins bind preferentially to the pyrimidine-rich strand of yeast and SV40 origins of replication, respectively (31). The significance, if any, of the interactions of *Tetrahymena* ssT-TIBF with T-rich oligonucleotides remains to be determined.

ssA-TIBF

Of the DNA-binding proteins categorized in this study, ssA-TIBF is the best candidate for a physiologically relevant factor. ssA-TIBF selectively binds to single-stranded DNA corresponding to the A-rich strand of the Type I repeat. This selectivity is indicated by the sequence-specific competition observed in binding assays and by the ~five-fold greater affinity ssA-TIBF exhibits for the A-rich strand of the Type I repeat relative to duplex Type I sequences (56). Since many single-stranded nucleic acid binding proteins bind both DNA and RNA, several oligoribonucleotides were also tested and shown to be poor competitors for ssA-TIBF binding (56).

In contrast to ds-TIBF, ssA-TIBF interacts most strongly with sequences at the ends of the conserved Type I repeat rather than the central AT-rich region. Under the binding conditions employed here, we were unable to detect a significant decrease in the affinity of ssA-

TIBF for oligonucleotides in which either one or five A residues were deleted from the central region. This was not unexpected, since the difference in the replication rates of the wildtype and *rmm* mutant rDNAs bearing a single A deletion is apparently $\leq 10\%$ in vegetatively dividing cells (35). However, in more detailed equilibrium binding studies, ssA-TIBF can discriminate between two alleles of the rDNA, C3 and B, that exhibit differential replication properties *in vivo* (See Paper 2).

These results suggest a possible role for ssA-TIBF in the control of rDNA replication. The replication defects observed in rDNA molecules bearing mutations in Type I repeats are consistent with a role for the Type I sequence in regulating the rate of replication initiation at the rDNA origin. Since ssA-TIBF binds single-stranded DNA preferentially, specific complexes between ssA-TIBF and Type I sequences *in vivo* must form after the generation of single-stranded regions. If initiation of DNA synthesis occurs within previously unwound regions in *Tetrahymena*, as it appears to do in some other eukaryotes (1, 25), ssA-TIBF could be involved in the initial recognition of the origin. Alternatively, ssA-TIBF could recognize the Type I sequence in duplex DNA (albeit with reduced affinity) and promote destabilization of the helix in the origin region. This possibility is currently under investigation. Finally, ssA-TIBF might function in events that occur after recognition of the origin by an unidentified initiator protein(s).

Sequence-specific single-stranded DNA-binding proteins

An increasing number of single-stranded DNA-binding proteins have been reported to interact specifically with DNA sequences involved in transcriptional regulation and recombination (7, 19, 21, 30, 32, 46, 48, 66). Moreover, several single-stranded DNA binding proteins that recognize sequences within chromosomal replication origin regions in eukaryotes have been identified recently. These include proteins that interact with the T-rich strand of ARS core consensus sequence in yeast (29, 33, 34, 49) and a mammalian single-stranded DNA binding protein that recognizes a repeated purine-rich sequence element, termed PUR, associated with several origins of replication and gene flanking regions in several of eukaryotes (5). However, the preferential binding of ssA-TIBF to an A-rich sequence element distinguishes it from all of the sequence-specific single-stranded DNA binding proteins noted above.

Defining the functional origin of replication *Tetrahymena* rDNA

Chromosomal DNA replication in many eukaryotes appears to initiate throughout broad regions rather than at the defined sites characteristic of *S. cerevisiae* (4). Initiation zones have been localized within regions of several hundred to several thousand base pairs in mammalian chromosomes (22, 62, 63) as well as in *Drosophila* (27, 52), and *S. pombe* (73). Although no specific DNA sequence or structural element within any of these potential origins has been directly implicated in the initiation of chromosomal DNA replication, initiation

zones in these organisms share modular elements that may contribute to a functional origin (4, 22, 36). These modular elements include DNA unwinding elements (DUEs), pyrimidine tracts that may serve as strong DNA polymerase-primase start sites, scaffold associated regions, and transcriptional regulatory sequences. Since it has been possible to map sites of replication initiation in *S. cerevisiae* with much greater precision than in the initiation zones mentioned above, it appears that *S. cerevisiae* origins are relatively compact. Although the ARS core consensus sequence is clearly an essential component of a replication origin in *S. cerevisiae*, it appears, however, that a fully functional chromosomal origin contains multiple functional elements (39, 65), including at least one or two of the modular elements found in the less well-defined origin regions of other eukaryotes.

In *Tetrahymena*, a 1.9 kb restriction fragment comprising the entire 5'NTS has been shown to promote autonomous replication of plasmids in the macronucleus (71), but it has not yet been possible to further delimit the region essential for rDNA origin function. Indeed, the identification of *rmm* mutations in two copies of the Type I repeat separated by more than 600 bp argues that cis-acting control sequences for rDNA replication may encompass an unexpectedly large region. The 5'NTS of *Tetrahymena* rDNA comprises a large (1.5kb) DUE and several other modular elements, including regions of bent DNA, scaffold attachment sites, and ARS-like sequences clustered in the vicinity of the Type I repeat sequences (36). Determination of the possible contribution of each of these elements to origin function will

require a combination of genetic, biochemical and molecular approaches.

A possible role for the Type I repeats in both transcription and replication

An interdependence of transcription and replication control has long been recognized in prokaryotes, and has been demonstrated recently in eukaryotes (reviewed in 14, 28). Cis-acting transcriptional control elements are found in association with replication origins in many eukaryotic DNA viruses, as well as in yeast (39, 65), *Drosophila* (41), CHO cells (8, 22) and *Tetrahymena* (35, 40). Furthermore, DNA-binding proteins that function as transcriptional activators or repressors have been shown to play a role in DNA replication. For example, transcription factors stimulate the replication of SV40, EBV, and polyomavirus through a mechanism that requires the transcriptional activation domains of the cognate factors (14, 28). A similar mechanism appears to be involved in the activation of yeast ARS sequences by transcription factor binding sites located near the ARS core consensus sequence (39).

We propose that the Type I repeats also play a role in the control of both transcription and replication of the rDNA in *Tetrahymena* (35, 40). There are four copies of the Type I repeat within the 5'NTS of the rDNA in *T. thermophila* (12). Two are located close to the rRNA transcription initiation site, and two lie within the nuclease hypersensitive Domains 1 and 2 (Fig. 1). The Type I repeat located closest to the rRNA transcription initiation site is essential for

promoter function *in vitro* (40). In view of the identification of the Type I repeat as a component of the transcriptional promoter, it is intriguing that several *rmm* mutations responsible for rDNA replication defects lie in upstream copies of the repeat (indicated by asterisks in Fig. 1C). The upstream repeats of the Type I element are reminiscent of the promoter element repeats in the rDNA spacers of *Xenopus* and several other eukaryotes(44). In wildtype rDNA, the essential promoter Type I repeat differs in 8 out of 33 nucleotides from the three upstream copies, which are identical in sequence (12). Nevertheless, ssA-TIBF binds this Type I repeat in the promoter region and the upstream Type I repeat in the replication origin region with equal affinity (54). The binding properties of ssA-TIBF *in vitro*, therefore suggest that it could be involved in the regulation of both rDNA replication and transcription *in vivo*.

ACKNOWLEDGMENTS

We are grateful to R. Benbow for many valuable suggestions and critical reading of the manuscript. We thank E. Henderson for comments on the manuscript, members of the Larson and Henderson laboratories for helpful discussions, and Ron Pearlman for providing the *T. thermophila* Type II repeat sequence oligonucleotides. This work was supported by NIH grant GM41899 (D.D.L.) and funds from the Iowa State University Biotechnology Council.

REFERENCES CITED

1. **Baldari, C.T., F. Amaldi, and M. Buongiorno-Nardelli.** 1978. Electron microscopic analysis of replicating DNA of sea urchin embryos. *Cell* **15**:1095-1107.
2. **Baur, C.-P. and R. Knippers.** 1988. Protein-induced bending of the Simian Virus 40 origin of replication. *J. Mol. Biol.* **203**: 1009-1019.
3. **Bell, S.P. and B. Stillman.** 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**:128-134.
4. **Benbow, R.M., J. Zhao, and D.D. Larson.** 1992. On the nature of origins of DNA replication in eukaryotes. *BioEssays* **10**:661-670.
5. **Bergemann, A.D. and E.M. Johnson.** 1992. The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. *Mol. Cell. Biol.* **12**:1257-1265.
6. **Bradford, M.M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72**:248-254.

7. **Brunel, F., P.M. Alzari, P. Ferrara, and M.M. Zakin.** 1991. Cloning and sequencing of PYBP, a pyrimidine-rich specific single strand DNA-binding protein. *Nucleic Acids Res.* **19**:5237-5245.
8. **Caddle, M.S., R.H. Lusier, and N.H. Heintz.** 1990. Intramolecular DNA triplexes, bent DNA and DNA unwinding elements in the initiation region of an amplified dihydrofolate reductase replicon. *J. Mol. Biol.* **211**:19-33.
9. **Campbell, J.L. and C.S. Newlon.** 1991. Chromosomal DNA replication, p. 41-146. *In* J.R. Broach, J.R. Pringle, and E.W. Jones (ed.), *The Molecular and Cellular Biology of the Yeast *Saccharomyces*: Genome Dynamics, Protein Synthesis and Energetics*. Cold Spring Harbor Press, Cold Spring Harbor, New York.
10. **Cech, T.R.** 1986. Ribosomal RNA gene expression in *Tetrahymena*: Transcription and RNA splicing, p. 203-225. *In* J.G. Gall (ed.), *The molecular biology of ciliated protozoa*. Academic Press, Inc., Orlando.
11. **Cech, T.R. and S.L. Brehm.** 1981. Replication of the extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila*. *Nucleic Acids Res.* **9**:3531-3543.

12. **Challoner, P.B., A.A. Amin, R.E. Pearlman, and E.H. Blackburn.** 1985. Conserved arrangements of repeated DNA sequences in nontranscribed spacers of ciliate ribosomal RNA genes: evidence for molecular coevolution. *Nucleic Acids Res.* **13**:2661-2680.
13. **Delidakis, C. and F.C. Kafatos.** 1989. Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila*. *EMBO J.* **8**:891-901.
14. **DePamphilis, M.L.** 1988. Transcriptional elements as components of eukaryotic origins of DNA replication. *Cell* **52**:635-638.
15. **Deshpande, A.M. and C.S. Newlon.** 1992. The *ARS* consensus sequence is required for chromosomal origin function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**:4305-4313.
16. **Diffley, J.F.X. and J.H. Cocker.** 1992. Protein-DNA interactions at a yeast replication origin. *Nature* **357**:169-172.
17. **Fairman, M.P. and B. Stillman.** 1988. Cellular factors required for multiple stages of SV40 replication *in vitro*. *EMBO J.* **7**:1211-1218.
18. **Fangman, W.L. and B.J. Brewer.** 1991. Activation of replication origins within yeast chromosomes. *Annu. Rev. Cell Biol.* **7**:375-402.

19. **Finocchiaro, L.M.E., P. Amati, and G.C. Glikin.** 1991. Single strand binding protein specific for the polyoma early-coding strand of PEA1 (AP1) regulatory sequence. *Nucleic Acids Res.* **19**:4279-4287.
20. **Fried, M. and D.M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505-6525.
21. **Gallard, C. and F. Strauss.** 1990. Sequence-specific single-strand-binding protein for the Simian Virus 40 early promoter stimulates transcription in vitro. *J. Mol. Biol.* **215**:245-255.
22. **Gale, J.M., R.A. Tobey, and J.A. D'Anna.** 1992. Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells. *J. Mol. Biol.* **224**:343-358.
23. **Galli, I., S.M.M. Iguchi-Ariga, and H. Ariga.** 1992. The AT-rich tract of the SV40 ori core: negative synergism and specific recognition by single stranded and duplex DNA binding proteins. *Nucleic Acids Res.* **20**:3333-3339.
24. **Garner, M.M. and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**:3047-3060.

25. **Gaudette, M.F. and R.M. Benbow.** 1986. Replication forks are underrepresented in chromosomal DNA of *Xenopus laevis* embryos. *Proceedings of the National Academy of Sciences of the United States of America* **83**:5953-5957.
26. **Greider, C.W. and E.H. Blackburn.** 1985. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **43**:405-413.
27. **Heck, M.M.S. and A.C. Spradling.** 1990. Multiple replication origins are used during *Drosophila* chorion gene amplification. *J. Cell Biol.* **110**:903-914.
28. **Heintz, N.H., L. Dailey, D.P. Held, and N. Heintz.** 1992. Eukaryotic replication origins as promoters of bidirectional DNA synthesis. *Trends in Genetics* **8**:376-381.
29. **Hofmann, J.F.-X. and S.M. Gasser.** 1991. Identification and purification of a protein that binds the yeast *ARS* consensus sequence. *Cell* **64**:951-960.
30. **Jansen-Durr, P., M. Boshart, B. Lupp, A. Bosserhoff, R.W. Frank, and G. Schutz.** 1992. The rat poly pyrimidine tract binding protein (PTB) interacts with a single-stranded DNA motif in a liver-specific enhancer. *Nucleic Acids Res.* **20**:1243-1249.

31. **Kim, C., R.O. Snyder, and M.S. Wold.** 1992. Binding properties of replication protein A from human and yeast cells. *Mol. Cell. Biol.* **12**:3050-3059.
32. **Kolluri, R., T.A. Torrey, and A.J. Kinniburgh.** 1992. A CT promoter element binding protein: definition of a double-strand and a novel single-strand DNA binding motif. *Nucleic Acids Res.* **20**:111-116.
33. **Kuno, K., S. Kuno, K. Matsushima, and S. Murakami.** 1991. Evidence for the binding of at least two factors, including T-rich strand-binding factors(s) to the single-stranded *ARS1* sequence in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **230**:45-48.
34. **Kuno, K., S. Murakami, and S. Kuno.** 1990. Single-strand-binding factor(s) which interact with *ARS1* of *Saccharomyces cerevisiae*. *GENE* **95**:73-77.
35. **Larson, D.D., E.H. Blackburn, P.C. Yaeger, and E. Orlas.** 1986. Control of rDNA replication in *Tetrahymena* involves a *cis*-acting upstream repeat of a promoter element. *Cell* **47**:229-240.
36. **Larson, D.D., W.-L. Shaiu, and R.M. Benbow.** Unpublished data

37. **Li, J. and T.J. Kelly.** 1985. Simian virus 40 DNA replication in vitro: Specificity of initiation and evidence for bidirectional replication. *Mol. Cell. Biol.* **5**:1238-1246.
38. **Malkas, L.H. and E.F. Baril.** 1989. Sequence recognition protein for the 17-base-pair A+T-rich tract in the replication origin of simian virus 40 DNA. *Proc. Natl. Acad. Sci. USA* **86**:70-74.
39. **Marahrens, Y. and B. Stillman.** 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**:817-823.
40. **Miyahara, K., N. Hashimoto, T. Higashinakagawa, and R.E. Pearlman.** in Press. Common sequence elements are important for transcription and replication of the extrachromosomal rRNA genes of *Tetrahymena*. *Gene*
41. **Orr-Weaver, T.L.** 1991. *Drosophila* chorion genes: cracking the eggshell's secrets. *Bioessays* **13**:
42. **Palen, T.E. and T.R. Cech.** 1984. Chromatin structure at the replication origins and transcription initiation regions of the ribosomal RNA genes of *Tetrahymena*. *Cell* **36**:933-942.

43. **Pan, W.C., E. Orias, M. Flacks, and E.H. Blackburn.** 1982. Allele-specific, selective amplification of a ribosomal RNA gene in *Tetrahymena thermophila*. *Cell* **28**:595-604.
44. **Reeder, R.H.** 1990. rRNA synthesis in the nucleolus. *Trends in Genet.* **6**:390-395.
45. **Rivier, D.H. and J. Rine.** 1992. An origin of DNA replication and a transcription silencer require a common element. *Science* **256**:659-663.
46. **Roller, R.J., A.L. McCormick, and B. Roizman.** 1989. Cellular proteins specifically bind single- and double-stranded DNA and RNA from the initiation site of a transcript that crosses the origin of DNA replication of herpes simplex virus 1. *Proc. Natl. Acad. Sci. USA* **86**:6518-6522.
47. **Sambrook, J., E.F. Fritsch, and T. Maniatis.** 1989. *Molecular cloning: A laboratory manual*. Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
48. **Santoro, I.M., T.-M. Yi, and K. Walsh.** 1991. Identification of single-stranded-DNA-binding proteins that interact with muscle gene elements. *Mol. Cell. Biol.* **11**:1944-1953.

49. **Schmidt, A.M.A., S.U. Herterich, and G. Krauss.** 1991. A single-stranded DNA binding protein from *S. cerevisiae* specifically recognizes the T-rich strand of the core sequence of ARS elements and discriminates against mutant sequences. *EMBO J.* **10**:981-985.
50. **Schulman, I.G., R.G. Cook, R. Richman, and C.D. Allis.** 1987. *Tetrahymena* contain two distinct and unusual high mobility group (HMG)-like proteins. *J. Cell Biol.* **104**:1485-1494.
51. **Shiau, W.L., P. Yaeger, M.F. Chau, E.H. Blackburn, E. Orias, and D. Larson.** Unpublished data.
52. **Shinomiya, T. and S. Ina.** 1991. Analysis of chromosomal replicons in early embryos of *Drosophila melanogaster* by two-dimensional gel electrophoresis. *Nucleic Acids Res.* **19**:3935-3941.
53. **Solomon, M.J., R. Strauss, and A. Varshavsky.** 1986. A mammalian high mobility group protein recognizes any stretch of six A-T base pairs in duplex DNA. *Proc. Natl. Acad. Sci. U.S.A.* **83**:
54. **Stahl, H., P. Droge, and R. Knippers.** 1986. DNA helicase activity of SV40 large tumor antigen. *EMBO J.* **5**:1939-1944.

55. **Stillman, B.W. and Y. Gluzman.** 1985. Replication and supercoiling of simian virus 40 DNA in cell extracts from human cells. *Mol. Cell. Biol.* **5**:2051-2060.
56. **Tjian, R.** 1978. The binding site on SV40 DNA for a T antigen-related protein. *Cell* **13**:165-179.
57. **Traut, W. and E. Fanning.** 1988. Sequence-specific interactions between a cellular DNA-binding protein and the Simian Virus 40 origin of DNA replication. *Mol. Cell. Biol.* **8**:903-911.
58. **Tsao, N.N.G. and R.E. Pearlman.** 1982. DNA binding proteins from *Tetrahymena thermophila*. *Can. J. Biochem.* **60**:398-407.
59. **Tullius, T.D. and B.A. Dombroski.** 1986. Hydroxyl radical "footprinting": High-resolution information about DNA-protein contacts and application to lambda repressor and *Cro* protein. *Proc. Natl. Acad. Sci. U.S.A* **83**:5469-5473.
60. **Umek, R.M. and D. Kowalski.** 1988. The ease of DNA unwinding as a determinant of initiation at yeast replication origins. *Cell* **52**:559-567.

61. **Vassilev, L. and E.M. Johnson.** 1990. An initiation zone of chromosomal DNA replication located upstream of the c-myc gene in proliferating HeLa cells. *Mol. Cell. Biol.* **10**:4899-4904.

62. **Vaughn, J.P., P.A. Dijkwel, L.H.F. Mullenders, and J.L. Hamlin.** 1990. Replication forks are associated with the nuclear matrix. *Nucleic Acids Res.* **18**:1965-1969.

63. **Walker, S.S., S.C. Francesconi, and S. Eisenberg.** 1990. A DNA replication enhancer in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **87**:4665-4669.

64. **Walker, S.S., A.K. Malik, and S. Eisenberg.** 1991. Analysis of the interactions of functional domains of a nuclear origin of replication from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **19**:6255-6262.

65. **Wilkison, W.O., H.Y. Min, K.P. Claffey, B.L. Satterberg, and B.M. Spiegelman.** 1990. Control of the adipin gene is adipocyte differentiation. *J. Biol. Chem.* **265**:477-482.

66. **Wobbe, C.R., L. Weissbach, J.A. Borowiec, F.B. Dean, Y. Murakami, P. Bullock, and J. Hurwitz.** 1987. Replication of Simian Virus 40 origin-containing DNA *in vitro* with purified proteins. *Proc. Natl. Acad. Sci. U.S.A.* **84**:1834-1838.

67. **Wold, M.S. and T.J. Kelly.** 1988. Purification and characterization of replication protein A, a cellular protein required for *in vitro* replication of simian virus 40 DNA. Proc. Natl. Acad. Sci. USA **85**:2523-2527.
68. **Yaeger, P.C., E. Orias, W.-L. Shaiu, D.D. Larson, and E.H. Blackburn.** 1989. The replication advantage of a free linear rRNA gene is restored by somatic recombination in *Tetrahymena thermophila*. Mol. Cell. Biol. **9**:452-460.
69. **Yao, M.-C.** 1986. Amplification of ribosomal RNA genes, p. 179-201. In J.G. Gall (ed.), The molecular biology of ciliated protozoa. Academic Press, Inc., Orlando, Florida.
70. **Yu, G.-L. and E.H. Blackburn.** 1989. Transformation of *Tetrahymena thermophila* with a mutated circular ribosomal DNA plasmid vector. Proc. Natl. Acad. Sci. U S A **86**:8487-91.
71. **Yu, G.L. and E.H. Blackburn.** 1990. Amplification of tandemly repeated origin control sequences confers a replication advantage on rDNA replicons in *Tetrahymena thermophila*. Mol. Cell. Biol. **10**:2070-80.

72. Zhu, J., C. Brun, K. H., Y. M., and J.A. Huberman. 1992. Identification and characterization of a complex chromosomal replication origin in *Schizosaccharomyces pombe*. *Chromosoma* **102** (Supplement):S7-S16.

PAPER 2. A DNA-BINDING PROTEIN INTERACTS WITH BOTH THE
REPLICATION ORIGIN AND THE TRANSCRIPTIONAL PROMOTER OF
THE rRNA GENES IN *TETRAHYMENA THERMOPHILA*

**A DNA-Binding Protein Interacts with both the Replication
Origin and the Transcriptional Promoter
of the rRNA Genes in *Tetrahymena thermophila***

Angela R. Umthun, Drena D. Larson*

Department of Zoology and Genetics
Signal Transduction Training Group
Molecular, Cellular and Developmental Biology Program
Interdepartmental Genetics Program

Room 2114 Molecular Biology Building
Iowa State University, Ames, IA 50011

* Corresponding author.

Phone: (515)294-1112

FAX: (515)294-2876

Email: ddlarson@molebio.iastate.edu

ABSTRACT

A single-stranded DNA-binding protein, ssA-TIBF, that specifically interacts with the A-rich strand of the Type I repeat, a cis-acting replication control element, in the 5' non-transcribed spacer region of the rRNA genes (rDNA) in *Tetrahymena thermophila* has been described (39). ssA-TIBF demonstrates strict sequence requirements for binding: Elimination of either end of the Type I repeat dramatically decreases binding of ssA-TIBFs. Here we report the purification of ssA-TIBF and further characterization of the ssA-TIBF binding site.

Two natural rDNA variants, C3 and B, differ in replication efficiency and in the sequences flanking the 3' end of one of the copies of the Type I repeat (28). ssA-TIBF binds preferentially to the 3' extended forms of the C3 rDNA Type I repeat. In addition, ssA-TIBF binds to the Type I repeat immediately adjacent to the transcriptional start site, which has been shown to be an essential element of the promoter *in vitro* (33). The affinity of ssA-TIBF for the extended form of this repeat equals the affinity for the upstream Type I repeats, even though this repeat differs from the upstream repeats at 8/33 positions. These results suggest that ssA-TIBF may be a transcription factor that is also involved in the regulation of rDNA replication.

INTRODUCTION

The difficulty of defining precise origins of replication in eukaryotes has been one key problem in determining the mechanism of replication initiation. Specific eukaryotic cis-acting sequences that function as sites of DNA replication initiation have been elusive. In yeast, a subset of the currently defined ARS (Autonomously Replicating Sequences) elements function as chromosomal origins of replication (5). The essential ARS element consists of an 11 bp core consensus sequence (ACS) flanked by an AT-rich region containing several functional sites, including binding sites for transcription factors. Two types of protein-DNA complexes have been shown to interact with the ARS core consensus sequence. One of these specifically interacts with the T-rich single-stranded form of the ARS core consensus sequence (23, 27, 37). A second factor consists of 6 polypeptides and forms sequence-specific interactions with the ARS element in double-stranded DNA, which includes sequences flanking the element (2). Interactions of this factor with the ARS are element corroborated by similar footprints *in vivo* (15). However, the specific roles of these factors, and whether they are involved in the initiation of DNA replication or at a latter stage, is not known.

An essential component of the ARS element involves transcription factor binding sites (32). Several cases of transcription factors being involved in DNA replication have been described. Binding sites for transcription factors localize with many origins and zones of DNA

replication in viral (14), yeast (16), and mammalian systems (21, 24). In several viral systems, transcription factor binding sites have been shown to be required for and to stimulate DNA replication (13, 14, 19, 20, 22, 30, 31). Furthermore, the binding of transcription factors with specific transcriptional activation domains is necessary for the stimulatory effects the activation of both viral replication origins and yeast ARSs on plasmids (3, 10, 18, 32). Thus transcription factors play an important role in DNA replication.

Several cis-acting replication control elements have been identified in the rRNA genes (rDNA) of *Tetrahymena thermophila*. In *Tetrahymena*, the rDNA is a 21kb palindromic extrachromosomal unit that is amplified to 10,000 copies in the macronucleus, (reviewed in (42)). An origin of replication has been mapped by electron microscopy to the 5' non-transcribed spacer (5' NTS) of the rDNA (7). Yu and Blackburn have shown that a 1.9 kb restriction fragment containing the region can promote autonomous replication of plasmids in macronuclei of *Tetrahymena* cells (43). This gives functional evidence that the 5' NTS contains sequences essential for the initiation of rDNA replication. Within the 5' NTS are three types of conserved repeated elements, designated Type I, II, and III repeats. The sequence and organization of these elements within the 5' NTS are conserved in three ciliate species, *Tetrahymena thermophila*, *Tetrahymena pyriformis*, and *Glaucoma chattoni* (8, 34). The Type I repeat is a genetically-defined control element for rDNA replication: specific mutations within the 33 nt repeat cause rDNA replication defects *in vivo* (28, 41). There are

four copies of the Type I repeat within the 5' NTS region (see Fig. 1) (8). The Type I repeat nearest the transcriptional start site has been shown to be an essential element of the promoter *in vitro* (33). The three upstream repeats of the Type I promoter element are not essential for transcription *in vitro*, but could play a role *in vivo*.

Replication differences between two natural variants of the rDNA found in inbred strains B and C3 were utilized to identify cis-acting replication control elements. When cells homozygous for B or C3 rDNA are crossed, the C3 rDNA gradually predominates in the macronucleus after several generations in the dividing progeny cells (28). Mutations which reversed the replication preference were isolated in C3 cells. Molecular and genetic analysis revealed that three of the mutants differed from wildtype by a single base pair deletion in one of the upstream Type I repeats (see Fig. 1). C3 and B rDNA differ primarily by a 42 base pair deletion that removes sequences immediately adjacent to the Type I repeat in Domain 2 (see Fig. 1) (28, 41). Taken together these results show that the Type I repeat is involved in the regulation of rDNA replication in *Tetrahymena*.

We previously characterized DNA-binding proteins that interact with the Type I repeat (39). One of these, ds-Type I Binding Factor (ds-TIBF) interacts with the duplex form of the repeat and appears to recognize the AT-rich central region. ds-TIBF also has an affinity for nonspecific AT-rich competitors suggesting that ds-TIBF recognition of the Type I repeat is promiscuous. In contrast, ssA-Type I Binding Factor (ssA-TIBF) specifically binds to Type I sequences because

Fig. 1. Structure of the *Tetrahymena thermophila* macronuclear rDNA.

The top line shows a map of the 21 kb palindromic rDNA. Arrows indicate pre-rRNA transcription units; solid boxes, coding regions for the 17S, 5.8S and 26S rRNAs; thin lines, 5' and 3' non-transcribed spacer regions; hatched lines, telomeric sequences (28, 29). The bottom lines illustrate the organization of repeated sequence elements in the 5'NTS of C3 and B rDNA. Domains I and 2 refer to hypersensitive domains that correspond to an ~400 bp imperfect duplication (34). Proposed replication origin(s) (6, 7, 35) are shown above Domains 1 and 2. Members of the Type I, II and III repeats (8) conserved repeated elements are indicated. A 42 bp deletion causes the Type III repeat to overlap with the Type I repeat in Domain 2. Asterisks denote Type I elements known to be altered in mutants defective in rDNA replication (28, 41).

21 kb rDNA

deletion of either border sequence to the AT-rich center greatly reduces ssA-TIBF binding. To test whether ssA-TIBF binding could be influenced by sequences outside the conserved Type I repeat element we have further investigated of the sequence requirements for ssA-TIBF binding.

Using gel mobility shift assays we have further characterized the ssA-TIBF binding site. In addition to the conserved Type I repeat element, ssA-TIBF also interacts with sequences flanking the 3' end in two of the upstream repeats. These sequences are deleted in B rDNA. Competition experiments comparing the affinity of ssA-TIBF for extended versions of the Type I repeat from B versus C3 rDNA, reveal a preference for the C3 Type I repeat. Thus, the affinity of ssA-TIBF for C3 versus B rDNA Type I sequences mimics the replication hierarchy *in vivo*. Moreover, although the Type I repeat in the rDNA promoter region differs slightly from the upstream versions, competition assays indicate that the affinity of ssA-TIBF for these sequences is identical.

MATERIALS AND METHODS

Cell growth and S100 extract preparation

S100 extracts were prepared from *Tetrahymena* cell cultures essentially as previously described (39). *T. thermophila* strain C3V cells were grown at 30°C to mid-log phase (2.5×10^5 cells/ml) in 2% PPYS (2% proteose peptone, 0.2% yeast extract, 0.003% sequestrine). Cells were collected (centrifugation at $2600 \times g$, 7 min.) and washed twice in chilled 10 mM Tris-HCl, pH 7.5. The final cell pellet was resuspended in 5X the pellet volume of HMG (20 mM HEPES, pH 7.9, 1mM MgCl₂, 10% glycerol). Cells were lysed by addition of NP40 to 0.2% final concentration for 30 min at 4°C while stirring. The whole cell extract was centrifuged at $100,000 \times g$ for 60 min. in a Beckman TLA 100.3 fixed angle rotor. Aliquots of the supernatant were immediately frozen in liquid nitrogen, and stored at -80°C.

Column chromatography

ssA-TIBF was purified by 4 successive chromatographic steps. The ssA-TIBF binding activity was quantitated by gel mobility shift assays and protein concentrations were determined by Bradford protein assays. Conductivity readings were used to determine the NaCl concentration in each fraction. 40 ml of *T. thermophila* S100 extracts (3.7 mg/ml protein) was applied to a 80 ml (Whatman P11, 2.5 x 16.5 cm) phosphocellulose column equilibrated with HMG/0.1 (20 mM HEPES, pH 7.9, 1mM MgCl₂, 10% glycerol, 0.02% NP40, 1mM DTT, 0.1 mM PMSF, 0.01

mM pepstatin, 0.01 mM leupeptin, 0.1 mM NaCl). The column was washed with 300 ml HMG/0.1 and bound proteins were eluted with a step gradient (200 ml/step) of NaCl (0.15M, 0.2M, and 0.4M) in HMG. Peak ssA-TIBF containing fractions from two phosphocellulose columns were pooled (120 ml; 0.4M NaCl) and loaded onto a preppacked 5ml (1.5 x 3.7 cm) econo pac blue column (Biorad) equilibrated with HMG/0.4. The column was washed with 100 ml HMG/0.4 and bound proteins were eluted with a step gradient of NaCl (0.8M, 1M, 1.5M and 2M) in HMG. Half of the fractions containing ssA-TIBF binding activity were pooled (45 ml) and concentrated (Centricon Centriprep concentrators) to 2.8 ml. 11 ml of HMG/0 was added to the concentrated protein to dilute the NaCl to 0.2M. The diluted sample was once again concentrated for a final volume of 4.5 ml. This was loaded onto a HMG/0.2 equilibrated 6ml (1.2 x 8 cm) sequence-specific DNA-affinity column. This column consisted of ~70-95 ug of the Type I oligonucleotide 5'GGCAAAAAAAAAACAAAAATAGTAAACCTTCCGAAC3' (Nucleic Acid Facility, Iowa State University) covalently linked to a CNBr-activated sepharose matrix (Pharmacia) as previously described by Kadonaga and Tjian (26). The column was wash with 20 ml HMG/0.2 and bound protein was eluted with a step gradient of NaCl (0.3M, 0.5M, and 1M) in HMG. Peak ssA-TIBF containing fractions eluting at 0.5M NaCl were pooled, diluted and desalted as before and reapplied to the Type I DNA affinity column. Proteins were eluted from the second DNA-affinity column with the same step NaCl gradient used for the first DNA affinity column.

Gel filtration chromatography, as described previously (39), was used as an alternative procedure to separate Type I binding factors. 1 ml of S100 extract (3.5 mg/ml) was applied to a Sephacryl S-300 column (Pharmacia, 1.5 x 88 cm) equilibrated with TEG/0.2 (TEG with 0.2 M NaCl). Proteins were eluted with 200 ml of TEG/0.2 and 1 ml fractions were collected.

Gel mobility shift assays

Synthetic oligonucleotide substrates for gel mobility shift assays were purchased from the Nucleic Acid Facility at Iowa State University. The oligonucleotides were 5' end labeled using T4 polynucleotide kinase (Biorad) and purified using Micro Select D spin columns (5' Prime-3' Prime). Gel mobility shift experiments were performed essentially as described previously (39). 1 ng ³²P-labeled oligonucleotides plus 1 µg Poly dIdC were incubated with 5-12 µl of column fractions containing ssA-TIBF for 15 min on ice. When fractions beyond the phosphocellulose stage of purification (see column chromatography above) were assayed poly dIdC was left out. In competition assays, 1 µl of the appropriate concentration of unlabeled oligonucleotide in STE (100 mM NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA) was added. Reaction mixtures contained 17 mM HEPES, pH 7.9, 0.8 mM NaCl, 8.7% glycerol (v/v), 200 mM NaCl, 1 mM DTT, 0.087 mM PMSF in a total of 15 µl. Binding reactions were carried out in the following order: buffer, NaCl, poly dIdC, ³²P-labeled oligonucleotides, competitor oligonucleotides (when appropriate), with protein added last. Samples

were electrophoresed in 5% polyacrylamide gels in 0.6X Tris-borate-EDTA (36) at 200V (>50 mAmps) for 60 min. After electrophoresis, gels were dried and exposed to Kodak XAR-5 film. To determine the binding of oligonucleotides, the radioactivity in individual bands was quantitated in dried gels using a Molecular Dynamics PhosphorImager.

UV crosslinking

UV crosslinking was performed essentially as previously described (1). Binding reactions (15 μ l), were prepared essentially as described above for gel mobility shift assays. S100 *Tetrahymena* cellular extracts or partially purified gel filtration fractions were incubated with 32 P-labeled oligonucleotides corresponding to the 37mer extended Type I repeat (see figure 3 for sequence) for 15 min on ice. After DNA-protein complexes were allowed to form, the samples were irradiated using a 254 nm UV source <5 cm above the samples for 30 min on ice at 4°C. 5 μ l 4X SDS loading dye (36) was added and the reaction products were boiled for 5 min. After cooling, samples were loaded onto a 10% SDS-polyacrylamide gel and electrophoresed in Laemmli buffer (36) at 30 mA for 3-4 hrs. Marker lanes contained pre-stained molecular weight standards (Biorad). After electrophoresis the gel was soaked for 30 min in GM buffer (10% glycerol, 5% methanol), dried and exposed to Kodak XAR-5 film.

RESULTS

The DNA-binding activity of ssA-TIBF resides in a 95 kD polypeptide.

ssA-TIBF was purified as described in the Materials and Methods. Purification was achieved by the following chromatographic steps: phosphocellulose, econo pac blue, and two sequence-specific DNA affinity columns in which the 37 nt 3' end extended version of the Type I repeat was covalently linked to the CNBr-activated sepharose matrix. Gel mobility shift assays comparing the ssA-TIBF binding activity present in the pooled column fractions is shown in Fig. 2. Fractions from the final sequence-specific DNA affinity column produce a single shifted band representing the purified ssA-TIBF factor. The ssA-TIBF eluting from the econo pac blue column migrates in the position of the lowest band observed in gel mobility shift assays of crude extracts. Gel mobility shift competition assays show that the econo pac blue purified ssA-TIBF has the same binding affinity as the phosphocellulose purified ssA-TIBF (data not shown). This suggests that the altered mobility may be due to degradation of the partially purified factor, but does not affect the sequence specificity of the econo pac blue purified ssA-TIBF. An overall 870 fold purification of A-TIBF was achieved with this strategy as quantitated by gel mobility shift assays.

Since protein was not detectable on denaturing polyacrylamide gels by silver staining of the most highly purified fraction, the molecular weight of ssA-TIBF was determined by UV crosslinking. Fig. 3 shows

Fig. 2. Purification of ssA-TIBF.

Pooled fraction containing ssA-TIBF from consecutive columns were compared by gel mobility shift assays. ³²P-labeled 37 nt Type I repeat oligonucleotides was used as the binding substrate. Lane 1, S100 cellular *Tetrahymena* extracts; lane 2, pooled phosphocellulose fractions; lane 3, pooled econo pac blue fractions; lane 4, pooled fraction from the first sequence-specific DNA-affinity column; lane 5, one fraction from the second sequence-specific DNA-affinity column.

S100
Phosphocellulose
Econo pac blue
DNA affinity
DNA affinity

Fig. 3. UV crosslinking of ssA-TIBF with the 37 nt Type I repeat.

³²P-labeled 37 nt Type I repeat oligonucleotides were incubated with partially purified ssA-TIBF, were crosslinked by UV irradiation, and separated according to size by denaturing gel electrophoresis (see Materials and Methods). Complexes were allowed to form in the presence of unlabeled competitor oligonucleotides; lane 1, non-irradiated control; lane 2, no competitor; lane 3, ss dA₃₃, nonspecific competitor oligo dA₃₃; lane 4, ssA Type I, specific competitor single-stranded A-rich Type I repeat. The molecular weights of protein standards and their relative mobilities are indicated to the left of the gel.

that a polypeptide of approximately 105 kD was covalently linked to the ^{32}P -labeled Type I repeat probe when the protein-DNA complex was irradiated prior to electrophoresis (Fig. 3, lane 2). No band was found in reactions in which ssA-TIBF was not present (data not shown), or not irradiated (Fig. 3, lanes 1). The intensity of the band was unaffected by the presence of 50X molar excess of an unlabeled nonspecific competitor, oligo (dA)₃₃ (Fig. 3, lane 3). A 50 molar excess of unlabeled Type I repeat sequences included in the reaction mixture, however, reduced the intensity of the labeled band by greater than 80% (Fig. 3, lane 4). These results confirm that ssA-TIBF, specifically recognizes Type I repeat sequences in gel mobility shift competition assays (39). We conclude that the sequence-specific DNA binding activity of ssA-TIBF is associated with the single polypeptide observed in figure 3. The ssA-TIBF-associated polypeptide appears to be 95 kD, when the 10 kD contributed by the linked oligonucleotide is subtracted.

The native molecular weight of ssA-TIBF previously estimated on the basis of S-300 gel filtration chromatography was 250 kD (39). This suggests that ssA-TIBF may exist as a multimer in crude extracts.

ssA-TIBF binds to additional sequences on the 3' end of the Type I repeat.

ssA-TIBF binding to the Type I repeat exhibits sequence specificity (39). Deletion of sequences That form either border of the repeat greatly reduce binding to the Type I repeat. Sequences outside the Type I repeat, therefore, may also be involved in interactions of ssA-TIBF

with the Type I repeat. Gel mobility shift competition experiments confirm that sequences on the 3' end of the Type I repeat are indeed important (Fig. 4). Partially purified ssA-TIBF was allowed to bind to ³²P-radiolabeled conserved 33 nucleotide Type I repeats in the presence or absence of unlabeled competitor oligonucleotides. Competitors included the Type I repeat and oligonucleotides consisting of part of the Type I repeat plus additional sequences on the 3' end (Fig. 4A). The 3' extended competitor oligonucleotide consisted of the Type I repeat plus twelve 3' nucleotides derived from the sequences immediately adjacent to the Type I repeat in Domain 2 of the C3 rDNA. To maintain a consistent oligonucleotide length the seven Ts on the 5' end of the Type I repeat were deleted. An oligonucleotide deleting the seven Ts was also used as a competitor to ensure that they were not involved in the recognition of the Type I sequences by ssA-TIBF. The extended Type I competitor is able to completely deplete binding to the wild type 33mer Type I repeat (Fig. 4B). Quantitation of the results suggested that ssA-TIBF had a 2 fold higher affinity for the extended form of the Type I repeat. These results indicate that sequences 3' to the Type I repeat play an important role in the binding of ssA-TIBF.

Comparison of sequences surrounding the Type I repeats in the 5' NTS.

To investigate the significance of the flanking sequences we examined the four 5' NTS Type I repeats, designated Ia, Ib, Ic, and Id (Fig. 5A) in the context of their surrounding sequences (Fig. 5B). As

Fig. 4. The recognition sequence of ssA-TIBF includes 3' flanking sequences as well as the Type I repeat.

(A) Single-stranded Type I repeat oligonucleotide sequences and their binding affinities for ssA-TIBF. The oligonucleotides shown were used as competitors in the gel mobility shift assay in panel B: Type I repeat, the A-rich single-strand comprising the conserved 33 nucleotide Type I repeat; Extended Type I, 26 nt of the Type I repeat plus 12 nt that flank the 3' end of the repeat found within Domain 2 in C3 rDNA (see Fig. 1C); T₇ deletion, the conserved Type I repeat minus the 7Ts on the 5' end. Boxed nucleotide sequences indicate sequences derived from the conserved 33nt Type I repeat. The ability of the oligonucleotides to compete for ssA-TIBF binding is expressed relative to that of the input oligonucleotide, the 33 nt Type I repeat (100%). Values are from an average of 2 experiments.

(B) Gel mobility shift competition assay illustrating the affinity of ssA-TIBF for extended versions of the Type I repeat. ³²P-labeled 33 nt Type I repeat oligonucleotide was bound by partially purified ssA-TIBF (gel filtration column fraction) in the presence or absence of unlabeled competitor oligonucleotides (see Materials and Methods). Competitors used are indicated above each set of 3 lanes which show reactions containing 0, 10X, and 50X molar excess of competitor. The abbreviations are explained in panel A.

A

Synthetic Oligonucleotides:

A-rich strand of Type I repeat

		Competition ability (%)	Binding ability
Type I repeat (33 nt)	5' TTTTTTGGCAAAAAAAAAACAAAAATAGTAA 3'	100	+
Extended Type I (37 nt)	GGCAAAAAAAAAACAAAAATAGTAA ACCTTCCGAAC	200	+
T ₇ deletion (26 nt)	GGCAAAAAAAAAACAAAAATAGTAA	95-100	+
Nonspecific Competitor			
ss (dA) ₃₃	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	<20	-

B

Fig. 5. Additional sequences that enhance ssA-TIBF binding flank two Type I repeats identified as cis-acting replication control elements.

(A) Organization of repeated sequence elements of the C3 rDNA 5' NTS. Domain 2 is an imperfect repeat of Domain 1. Domain 1 and Domain 2 are nuclease hypersensitive domains. They represent ~400 bp of imperfect duplication. Black filled boxes designated Ia-d, four copies of the Type I repeat; Slashed boxes designated II, Type II repeats; stippled boxes designated III, Type III repeats which are topoisomerase I binding and cleavage sites. The wavy arrow indicates the start site for rRNA transcription. Astericks denote repeats shown to be altered in *rmm* mutants.

(B) Comparison of the four copies of the Type I repeats and flanking sequences in C3 rDNA. a-d correspond to the Type I repeats depicted in panel A. Boxed nucleotides are the previously described conserved Type I repeat sequences. 3' flanking nucleotides that are identical in b and c repeats are shown in shaded box.

C. Comparison of the Type I repeat (Ic) and 3' flanking sequences within Domain 2 of C3 and B rDNA. The boxed and shaded nucleotides are as described in panel B. Underlined nucleotides correspond to Type III repeat sequences that overlap the Type I repeat as a result of a 42 bp deletion in B rDNA relative to C3.

B) Type I repeat sequences in wildtype C3 rDNA

Ia: ATTAGATA TTTTTTTGGCAAAAAAAAAACAAAAATAGTAA AAAATCACTTTTTTTGAGAGTTGAAAA
 *Ib: ATTAGAAA TTTTTTTGGCAAAAAAAAAACAAAAATAGTAA ACCTTCCGAACITTT GCAACTTTTGAG
 *Ic: CAGAGTCT TTTTTTTGGCAAAAAAAAAACAAAAATAGTAA ACCTTCCGAACITTT TTGACTTTGAGA
 Id: TGAGAAA ATTCTTTGGCAAAAAAAAAATAAAATAATATCAG GGGGGTAAAAATGCATATTTAAGAAGG

C) Comparison of Type I repeat (Ib) and 3' flanking sequences in Domain 2

*C3(Ib): ATTAGAAA TTTTTTTGGCAAAAAAAAAACAAAAATAGTAA ACCTTCCGAACITTT GCAACTTTTGAG
 B(Ib): ATTAGAAA TTTTTTTGGCAAAAAAAAAACAAAAATAGTAA ACTTAGAAAAATTTTTGAAAAATGAAA

- = Conserved Type I Repeat
- = Conserved Type III Repeat
- = Conserved Adjacent Sequences
- * = Mutated Repeats in rmm Mutants

noted previously (28), the sequences that enhanced the affinity of ssA-TIBF for Type I repeats in the previous competition assay are conserved in the 3' flanking region of two of the Type I repeats (shaded region, Fig. 5B). Three individually isolated mutants of rDNA maturation and maintenance (C3-rmm 1, 3, and 4) have mutations which affect these Type I repeats (28). When the C3 and B rDNA sequences flanking the Domain 2 Type I repeat were compared (Fig. 5C), the B rDNA sequences evidently lack a portion of the ssA-TIBF binding site. This is due to the 42bp deletion which brings a Type III repeat immediately adjacent to the Type I element. However, it is important to point out that of the 15 nucleotides in the conserved 3' flanking sequences, 9 positions are still intact in B rDNA.

The binding affinity of ssA-TIBF for C3 and B extended Type I Repeats mimics the difference in replication efficiencies between C3 and B rDNA *in vivo*.

Can a difference in the affinity of ssA-TIBF for extended Type I elements explain the replication disadvantage B rDNA possesses in the presence C3 rDNA? A 53 base ³²P-labeled oligonucleotide consisting of the C3 Type I repeat and 3' flanking sequences competed for ssA-TIBF binding against unlabeled oligonucleotides reflecting deletions of both the C3 and B Type I repeats (Fig. 6). Gel mobility shift competition assays indicated that deleting the Type I repeat from the C3 oligonucleotides resulted in only a slight reduction in the ability of these sequences to bind (Fig. 6, lanes 1-9). This suggests that optimal

Fig. 6. The binding affinity of ssA-TIBF for C3 vs. B rDNA sequences *in vitro* parallel their replication efficiencies *in vivo*.

(A) Single-stranded Type I repeat oligonucleotide sequences and their relative ability to bind ssA-TIBF. The oligonucleotides shown were used as competitors in the gel mobility shift assay in panel B. The 53 nt extended Type I repeat consists of 26 nt of the Type I repeat plus 27 nt flanking the 3' end of the C3 rDNA Ic repeat; The 53 nt extended B Type I repeat consists of 26 nt of the Type I repeat plus 27 nt flanking the 3' end of the B rDNA Ic repeat. Boxed nucleotide sequences indicate sequences derived from the conserved 33 nt Type I repeat. Underlined nucleotide sequences indicate Type III repeat sequences. The ability of the oligonucleotides to compete for ssA-TIBF binding relative to the 53 nt input oligonucleotides is indicated. Values are from an average of 2 experiments.

(B) Gel mobility shift competition assay illustrating preferential binding of ssA-TIBF to Type I repeat 3' flanking sequences of C3 rDNA over B rDNA. P³²-labeled 53 nt Extended C3 Type I repeat oligonucleotides were bound by partially purified ssA-TIBF in the presence of competitor oligonucleotides (see Materials and Methods). Competitors used are indicated above each set of 3 lanes which show reactions containing 0, 10X, and 50X molar excess of competitor. The abbreviations are explained in panel A.

A

Synthetic Oligonucleotides:

A-rich strand of Type I repeat

Extended
C3 Type I53 nt 5' GGCAAAAAAAAAACAAAAATAGTAA ACCTTCGAACTTTTGCAACTTTGAGA 3'

100

+

35 nt AATAGTAA ACCTTCGAACTTTTGCAACTTTGAGA

100

+

27 nt ACCTTCGAACTTTTGCAACTTTGAGA

95

+

Extended
B Type I53 nt GGCAAAAAAAAAACAAAAATAGTAA ACTTAGAAAAATTTTGAAAAATGAAA

90-95

+

35 nt AATAGTAA ACTTAGAAAAATTTTGAAAAATGAAA

90-95

+

27 nt ACTTAGAAAAATTTTGAAAAATGAAA

40-50

-

Nonspecific Competitor

ss (dA)₃₃ AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

<20

-

B

binding occurs when both the conserved Type I repeat and 3' flanking sequences are present, however ssA-TIBF will interact strongly with a subset of the nucleotides in the Type I repeat. A similar reduction in competition ability was observed when either the 53 nt or 35 nt B rDNA sequences, containing all or part of the conserved Type I repeat, were used as competitors (Fig. 6, lanes 10-15). However, complete deletion of the Type I repeat caused ssA-TIBF to lose its affinity for the B rDNA sequences (Fig. 6, lanes 16-18). Identical results were obtained when the B oligonucleotide was the ^{32}P -labeled input (data not shown). Thus, ssA-TIBF can differentiate between C3 and B rDNA Type I repeat and 3' flanking sequences in Domain 2, with a slight preference for the C3 element.

ssA-TIBF specifically interacts with the Type I repeat that functions as an essential element of the promoter *in vitro*

The Type I repeat adjacent to the transcriptional start site is an essential promoter element in *in vitro* studies (33). The promoter element differs from upstream copies by 8 internal nucleotides and the 3' flanking sequences of the conserved 33 nt repeat sequence (see Fig. 5B). To test whether ssA-TIBF may bind to the Type I repeat promoter element we used the 37 nt extended version of the Type Ib repeat (see Fig. 5B) as the ^{32}P -labeled input and competed for ssA-TIBF binding with varying lengths of the extended Type Id promoter element (Fig. 7). The ability of oligonucleotides to compete for binding ssA-TIBF relative to the input is indicated to the right of the sequences in Fig.

Fig. 7. ssA-TIBF binds the promoter Type I repeat and upstream repeat sequences with an equal affinity.

(A) Single-stranded Type I repeat oligonucleotide sequences and their relative ability to bind ssA-TIBF. The oligonucleotides shown were used as competitors in the gel mobility shift assay in panel B. The extended Type I, 26 nt of the Type I repeat plus 12 nt that flank the 3' end of the repeat found within Domain 2 in C3 rDNA (Fig. 5); Extended promoter, 26 nt of the promoter Type I repeat plus 12 nt that flank the 3' end of the promoter repeat (Fig. 5); Truncated promoter, 15 nt of the promoter Type I repeat plus 12 nt that flank the 3' end of the promoter repeat. Boxed nucleotide sequences indicate sequences derived from the conserved 33 nt Type I repeat or the promoter Type I repeat. The ability of the oligonucleotides to compete for ssA-TIBF binding relative to the 37 nt extended Type I repeat input oligonucleotide is indicated. Values are from an average of 3-4 experiments.

(B) Gel mobility shift competition assay illustrating the affinity of ssA-TIBF for extended versions of both the upstream and promoter Type I repeat. P³²-labeled Extended Type I repeat oligonucleotides were bound by partially purified ssA-TIBF (gel filtration column fraction) in the presence of competitor oligonucleotides (see Materials and Methods). Competitors used are indicated above each set of 3 lanes which show reactions containing 0, 10X, and 50X molar excess of competitor. The abbreviations are explained in panel A.

A

Synthetic Oligonucleotides:		Competition ability	Binding ability
A-rich strand of Type I repeat			
Extended Type I (37 nt)	5' <u>GGCAAAAAAAAAA CAAAAATAGTAA</u> ACCTTCCGAAC 3'	100	+
Extended Promoter (38 nt)	<u>GGCAAAAAAAAAATAAAAAATAATATCAG</u> GGGGGTAAAAAT	100	+
Truncated Promoter (27 nt)	<u>TAAAAATAATATCAG</u> GGGGGTAAAAAT	65	-
Nonspecific Competitor			
ss (dA) ₃₃	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	<20	-

B

7A. Despite sequence variations, ssA-TIBF has a strong affinity for the promoter Type I repeat element. In this experiment, the 38 nt extended promoter element competed as effectively for ssA-TIBF as the extended Type Ib repeat (lanes 1-6). Deletion of a portion of the 33 nt conserved Type I region reduced the ability of the promoter sequences to compete for binding (lanes 7-9). Therefore, ssA-TIBF specifically interacts with both the promoter element and upstream copies of the Type I repeat.

DISCUSSION

Our results demonstrate that the ssA-TIBF binding site includes part of the Type I repeat and sequences flanking the 3' end of two Type I repeat elements within the 5' NTS. Both of these repeats have been implicated as cis-acting replication control elements *in vivo* (28, 41) (Yaeger, Shaiu, Larson, Orias unpublished data). These 3' flanking sequences are deleted in Domain 2 Type I repeats of the B strain rDNA, although 9 of the 15 nucleotide positions recognized by ssA-TIBF are intact. B rDNA is eventually lost in cells containing both B and C3 rDNA. This indicates that B rDNA molecules are less efficiently replicated (28).

We have shown by gel mobility shift assays that ssA-TIBF has a lower affinity for sequences flanking the 3' end of the Type I repeat of B rDNA. Although ssA-TIBF binds very strongly to the conserved 33 nt Type I element without the 3' sequences, it also appears, surprisingly, that ssA-TIBF interacts with sequences that flank the Type I repeat in the absence of the conserved element. This suggests that ssA-TIBF may recognize structural differences not immediately apparent in the primary sequence.

Differences in the affinity of ssA-TIBF for the 3' flanking sequences of C3 and B rDNAs are readily apparent. These difference are reflected in the 10% difference in the binding affinities observed with the longer forms of the repeat. In a physiological context these differences are sufficient to explain the replicative differences

observed between the rDNAs *in vivo*. A 10% difference in replication efficiency could account for the eventual dominance of C3 rDNA. From these data, we speculate that ssA-TIBF is involved in the physiological regulation of rDNA replication.

The Type I repeat nearest the transcriptional start site is an essential element of the promoter *in vitro* (33). The promoter element differs from the upstream copies by 8 nt within the 33 nt conserved Type I repeat and lacks the 3' sequences ssA-TIBF recognizes in upstream repeats. Despite these differences ssA-TIBF recognizes this promoter element in an extended 38mer oligonucleotide. The affinity of ssA-TIBF for the Type I repeat promoter element in gel mobility shift assays is equivalent to that of the extended upstream Type I repeat. This suggests that ssA-TIBF may be involved in rDNA transcription as well as replication.

Transcription factor binding sites are required for or stimulate replication of many mammalian viral genomes (reviewed in (14)). The presence of auxillary sequences containing binding sites for several transcription factors enhances simian virus 40 DNA replication both *in vivo* and *in vitro* up to 100 fold (13, 19, 22, 30, 31). The enhancement of replication is not due to the proximity to transcriptionally active DNA, because the initiation of replication is not sensitive to α -amanitin (12). Nor is chromatin modification solely responsible since replication of naked DNA templates is stimulated by these auxillary sequences (19). However, this enhancement requires that the auxillary sequences be in close proximity to the core origin (9, 25, 30). Similar

template requirements have been shown to be necessary for polyomavirus DNA replication (20). These data suggest that interactions between transcription factors and the replicative machinery stabilize or increase the activity of the replication factors.

In some cases, stimulation of DNA replication by transcription factors has been shown to require the activation domains of the transcription factor. In both polyomavirus and SV40 DNA replication, replacement of wildtype transcription factor binding sites by binding sites for other transcription factors stimulated DNA replication and inhibited repression during chromatin assembly (3, 10, 18).

In yeast the transcription factor ARS binding factor 1 (ABF1) may be important for ARS function. ABF1 binding sites occur at only a subset of origins, but the function of these origins in plasmid maintenance is reduced when these binding sites are interrupted (4, 16, 17, 38, 40). The role of ABF1 at the chromosome origin ARS1 has been characterized (32). Binding of ABF1 is required for plasmid maintenance. The binding site and binding of another transcription factor that also interacts with certain ARS elements, RAP1, are, however, able to substitute for ABF1. Substitution of the binding site for ABF1 with a binding site for a protein lacking the activation domain, contribute to the loss of the ABF1 function.

Binding sites for transcription factors have been localized in regions containing potential origins of replication in mammals. Transcriptional enhancer sequences and an origin of replication have been shown to localize upstream of the human *c-myc* gene (24). Several

transcription factors binding sites are located adjacent to the origin of bidirectional replication of the *dhfr* locus of Chinese Hamster Ovary Cells (21). These binding sites include Oct1, AP1, and RIP60. RIP60 purifies with a DNA helicase activity RIP100, suggesting that these factors may be involved in DNA unwinding events preceding DNA replication (11).

REFERENCES CITED

1. **Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl.** 1989. Current protocols in molecular biology. Greene publishing associates and Wiley-Interscience, New York.
2. **Bell, S.P. and B. Stillman.** 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**:128-134.
3. **Bennett-Cook, E.R. and J.A. Hassell.** 1991. Activation of polyomavirus DNA replication by yeast GAL4 is dependent on its transcriptional activation domains. *EMBO J.* **10**:959-969.
4. **Buchman, A.R. and R.D. Kornberg.** 1990. A yeast ARS-binding protein activates transcription synergistically in combination with other weak activating factors. *Mol. Cell. Biol.* **10**:887-897.
5. **Campbell, J.L. and C.S. Newlon.** 1991. Chromosomal DNA replication, p. 41-146. In J.R. Broach, J.R. Pringle, and E.W. Jones (ed.), *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis and Energetics*. Cold Spring Harbor Press, Cold Spring Harbor, New York.

6. **Cech, T.R.** 1986. Ribosomal RNA gene expression in *Tetrahymena*: Transcription and RNA splicing, p. 203-225. In J.G. Gall (ed.), *The molecular biology of ciliated protozoa*. Academic Press, Inc., Orlando.
7. **Cech, T.R. and S.L. Brehm.** 1981. Replication of the extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila*. *Nucleic Acids Res.* **9**:3531-3543.
8. **Challoner, P.B., A.A. Amin, R.E. Pearlman, and E.H. Blackburn.** 1985. Conserved arrangements of repeated DNA sequences in nontranscribed spacers of ciliate ribosomal RNA genes: evidence for molecular coevolution. *Nucleic Acids Res.* **13**:2661-2680.
9. **Chandrasekharappa, S.C. and K.N. Subramanian.** 1987. Effects of positions and orientation of the 72-base-pair-repeat transcriptional enhancer on replication from the simian virus 40 core origin. *J. Virol.* **61**:2973-2980.
10. **Cheng, L., J.L. Workman, R.E. Kingston, and T.J. Kelly.** 1992. Regulation of DNA replication in vitro by the transcriptional activation domain of GAL4-VP16. *Proc. Natl. Acad. Sci. USA* **89**:589-593.

11. **Dailey, L., M.S. Caddle, N. Heintz, and N.H. Heintz.** 1990. Purification of RIP60 and RIP100, mammalian proteins with origin-specific DNA-binding and ATP-dependent DNA helicase activities. *Mol. Cell. Biol.* **10**:6225-6235.
12. **Decker, R.S., M. Yamaguchi, R. Possenti, and M.L. DePamphilis.** 1986. Initiation of simian virus 40 DNA replication in vitro: aphidicolin causes accumulation of early-replicating intermediates and allows determination of the initial direction of DNA synthesis. *Mol. Cell. Biol.* **6**:3815-3825.
13. **DeLucia, A.L., S. Deb, K. Partin, and P. Tegtmeyer.** 1986. Functional interactions of the simian virus 40 core origin of replication with flanking regulatory sequences. *J. Virol.* **57**:138-144.
14. **DePamphilis, M.L.** 1988. Transcriptional elements as components of eukaryotic origins of DNA replication. *Cell* **52**:635-638.
15. **Diffley, J.F.X. and J.H. Cocker.** 1992. Protein-DNA interactions at a yeast replication origin. *Nature* **357**:169-172.
16. **Diffley, J.F.X. and B. Stillman.** 1988. Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc. Natl. Acad. Sci. USA* **85**:2120-2124.

17. **Eisenberg, S., C. Civalier, and B.K. Tye.** 1988. Specific interaction between a *Saccharomyces cerevisiae* protein and a DNA element associated with certain autonomously replicating sequences. *Proc. Natl. Acad. Sci. USA* **85**:743-746.
18. **Guo, Z.-S. and M.L. DePamphilis.** 1992. Specific transcription factors stimulate simian virus 40 and polyomavirus origins of DNA replication. *Mol. Cell. Biol.* **12**:2514-2524.
19. **Guo, Z.-S., C. Gutierrez, U. Heine, J.M. Sogo, and M.L. DePamphilis.** 1989. Origin auxiliary sequences can facilitate initiation of simian virus 40 DNA replication in vitro as they do in vivo. *Mol. Cell. Biol.* **9**:3593-3602.
20. **Hassell, J.A., W.J. Muller, and C.R. Mueller.** 1986. The dual role of the polyomavirus enhancer in transcription and DNA replication. *Cancer Cells (Cold Spring Harbor)* **4**:561-569.
21. **Held, P.G. and N.H. Heintz.** 1992. Eukaryotic replication origins. *Biophys. Acta.* **1130**:235-246.
22. **Hertz, G.Z. and J. Mertz.** 1986. Bidirectional promoter elements of simian virus 40 are required for efficient replication of the viral DNA. *Mol. Cell. Biol.* **6**:3513-3522.

23. **Hofmann, J.F.-X. and S.M. Gasser.** 1991. Identification and purification of a protein that binds the yeast ARS consensus sequence. *Cell* **64**:951-960.
24. **Iguchi-Ariga, S.M.M., T. Okazaki, T. Itani, M. Ogata, Y. Sato, and H. Ariga.** 1988. An initiation site of DNA replication with transcriptional enhancer activity present upstream of the c-myc gene. *The EMBO Journal* **7**:3135-3142.
25. **Innis, J.W. and W.A. Scott.** 1984. DNA replication and chromatin structure of simian virus insertion mutants. *Mol. Cell. Biol.* **4**:1499-1507.
26. **Kadonaga, J.T. and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* **83**:5889-5893.
27. **Kuno, K., S. Murakami, and S. Kuno.** 1990. Single-strand-binding factor(s) which interact with ARS1 of *Saccharomyces cerevisiae*. *GENE* **95**:73-77.
28. **Larson, D.D., E.H. Blackburn, P.C. Yaeger, and E. Orlas.** 1986. Control of rDNA replication in *Tetrahymena* involves a cis-acting upstream repeat of a promoter element. *Cell* **47**:229-240.

29. **Larson, D.D., E.A. Spangler, and E.H. Blackburn.** 1987. Dynamics of telomere length variation in *Tetrahymena thermophila*. *Cell* **50**:477-83.

30. **Lee-Chen, G.-J. and M. Woodworth-Gutai.** 1986. Simian virus 40 DNA replication: Functional organization of regulatory elements. *Mol. Cell. Biol.* **6**:3086-3093.

31. **Li, J.J., K.W.C. Peden, R.A.F. Dixon, and T. Kelly.** 1986. Functional organization of the simian virus 40 origin of DNA replication. *Mol. Cell. Biol.* **6**:1117-1128.

32. **Marahrens, Y. and B. Stillman.** 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**:817-823.

33. **Miyahara, K., N. Hashimoto, T. Higashinakagawa, and R.E. Pearlman.** in Press. Common sequence elements are important for transcription and replication of the extrachromosomal rRNA genes of *Tetrahymena*. *Gene*

34. **Niles, E.G. and K. Cunningham.** 1981. Structure of the *Tetrahymena pyriformis* rRNA gene: nucleotide sequence of the transcription initiation region. *J. Biol. Chem.* **256**:12857-12860.

35. **Palen, T.E. and T.R. Cech.** 1984. Chromatin structure at the replication origins and transcription initiation regions of the ribosomal RNA genes of *Tetrahymena*. *Cell* **36**:933-942.
36. **Sambrook, J., E.F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: A laboratory manual. Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
37. **Schmidt, A.M.A., S.U. Herterich, and G. Krauss.** 1991. A single-stranded DNA binding protein from *S. cerevisiae* specifically recognizes the T-rich strand of the core sequence of ARS elements and discriminates against mutant sequences. *EMBO J.* **10**:981-985.
38. **Shore, D., D.J. Stillman, A.H. Brand, and K.A. Nasmyth.** 1987. Identification of silencer binding proteins from yeast: possible roles in SIR control and DNA replication. *EMBO J.* **6**:461-467.
39. **Umthun, A.R., W.-L. Shaiu, Z.A. Sibenaller, and D.L. Larson.** submitted. A single-stranded DNA-binding protein recognizes a replication control element in the origin region of the rDNA in *Tetrahymena thermophila*. *Mol. Cell. Biol.*
40. **Walker, S.S., S.C. Francesconi, and S. Eisenberg.** 1990. A DNA replication enhancer in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **87**:4665-4669.

41. **Yaeger, P.C., E. Orias, W.-L. Shaiu, D.D. Larson, and E.H. Blackburn.** 1989. The replication advantage of a free linear rRNA gene is restored by somatic recombination in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **9**:452-460.
42. **Yao, M.-C.** 1986. Amplification of ribosomal RNA genes, p. 179-201. In J.G. Gall (ed.), *The molecular biology of ciliated protozoa*. Academic Press, Inc., Orlando, Florida.
43. **Yu, G.L. and E.H. Blackburn.** 1990. Amplification of tandemly repeated origin control sequences confers a replication advantage on rDNA replicons in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **10**:2070-2080.

GENERAL CONCLUSIONS

A complete understanding of how eukaryotic DNA replication is controlled requires an understanding of the DNA sequences and protein factors involved in the initiation event. To begin to understand the mechanisms controlling rDNA replication in *Tetrahymena*, I have characterized three DNA-binding factors that interact with a genetically-defined cis-acting replication control element, the Type I repeat. The three Type I binding factors (TIBFs) differ in their abilities to bind conformationally different Type I repeat sequences. One of these factors, ds-TIBF, interacts specifically with the AT-rich central region of the duplex form of the repeat. Two of the factors prefer a single-stranded Type I repeat conformation, each interacting with either the A-rich or the T-rich single-strand repeat sequence. The T-rich TIBF factor(s), ssT-TIBF, do not specifically recognize the Type I repeat since any T-rich sequence can compete for binding. However, the ssA-rich TIBF factor(s) specifically interacts with the sequences flanking the A-rich central region of the Type I repeat. Although these factors have variable affinities for the Type I repeat, it is possible that all three of these factors plays a physiological role in rDNA replication.

Although the duplex Type I binding factor interacts with poly (dAdT)₃₃ as well as the Type I repeat sequences, it demonstrates some sequence-specific interactions. Competitors with less than 6 central AT-base pairs are unable to compete for binding ds-TIBF. This includes

the Type II repeat sequences which also lie within the 5' NTS. The affinity of ds-TIBF for competitors consisting of strictly AT base pairs of 25-33 nt in length, such as oligo (dAdT)₃₃, suggests that AT-richness is important for binding. It is important to note that theoretically these competitors have multiple binding sites which may influence their competitiveness. However, competitors with longer stretches of AT base pairs can compete quite well, even if the stretches are interrupted by a CG base pair. ds-TIBF also exhibits a strong preference for the duplexed form of the Type I repeat since the single-strands are unable to compete for binding in gel mobility shift assays. Finally, nonspecific RNA polymers are not able to compete for ds-TIBF binding to the Type I repeat (data not shown) suggesting that ds-TIBF binding is DNA specific. These data suggest that ds-TIBF may be specifically recognizing the duplexed AT-rich central region of the Type I repeat.

In experiments with ssA-TIBF, we have shown that sequences flanking the 3' end of the Type I repeat are recognized by ssA-TIBF. Perhaps sequences specifically recognized by ds-TIBF are not present in the oligonucleotides reflecting the conserved Type I repeat. In the future, it will be important to examine the interactions of ds-TIBF with the Type I repeat in the context of its flanking sequences. It is also possible that ds-TIBF may interact more specifically with the Type I repeat *in vivo*. The specific recognition of the Type I repeat by ds-TIBF may involve protein-protein interactions that are not formed

in gel mobility assays. These possibilities will need to be addressed in future experiments.

There are precedent for the involvement of AT-rich binding proteins in DNA replication. Several proteins have been shown to recognize a 17-base-pair A/T-rich tract which is part of the minimal core origin necessary for SV40 DNA replication (42, 72). One of these factors was isolated from a multienzyme complex including the DNA polymerase α -primase complex (72). However, unlike ds-TIBF which prefers duplexed DNA, some of these proteins specifically bind both the single stranded and duplex forms of the A/T-rich tract.

ssT-TIBF(s), identified by gel mobility shift assays, binds to the T-rich single-stranded form of the Type I repeat. The interactions of ssT-TIBF with the Type I repeat appear to be nonspecific since any T-rich single-stranded oligonucleotide can compete for ssT-TIBF binding. Because of the lack of ssT-TIBF sequence-specific recognition of the Type I repeat we did not pursue further characterization of these factors.

A nonspecific single-stranded DNA binding protein that exhibits a preference for pyrimidine tracts is essential for SV40 DNA replication *in vitro* (62). RP-A (also known as RF-A or SSB in humans) has been isolate from human (58), *Crithidia* (13) and yeast (12). RP-A facilitates the unwinding event in SV40 DNA replication, however *E. coli* SSBs can substitute for this function. *E. coli* SSBs cannot substitute for the RP-A functional requirement for SV40 DNA replication (109). Therefore, it is thought that one of the functional

roles of RP-A may involve specific protein-protein interaction with the replication machinery.

Since, nonspecific single-stranded DNA binding proteins are essential for DNA replication in other systems, it will be important to further characterize ssT-TIBF and its role in rDNA replication. An *in vitro* system of DNA replication using *Tetrahymena* cellular extracts is currently being developed in Drena Larson's laboratory. Once an *in vitro* system of DNA replication is achieved, it will be possible to test the role of ssT-TIBF in rDNA replication.

In contrast to both the previously described Type I binding factors, the A-rich single-stranded Type I Binding Factor (ssA-TIBF) exhibits strict sequence requirements for recognition. ssA-TIBF cannot bind Type I oligonucleotides in which either end of the conserved repeat is deleted, suggesting that ssA-TIBF interacts with nucleotides in these regions. Further characterization of the sequences recognized by ssA-TIBF, revealed that ssA-TIBF recognizes specific nucleotides that normally flank the 3' end of a subset of Type I repeats. The two repeats flanked by these sequences have both been shown to be altered in C3-*rmm* mutants. A comparison of C3 and B rDNA sequences flanking the Type I repeat in Domain 2 reveals that a 42 base pair deletion in B rDNA removes the 3' flanking sequences recognized by ssA-TIBF. However, 9 of the 15 nucleotide positions conserved in the 3' flanking sequences are not altered by the deletion, suggesting that ssA-TIBF may still be able to recognize these sequences. When the affinity of ssA-TIBF for the C3 versus B 3' extended Type I repeats is compared, a

5-10% preference for the C3 sequences is revealed. This mimics the replicative capabilities of these rDNA when they compete for replication factors within a single macronucleus. Only a 10% difference in replication rates is enough to explain the eventual dominance of C3 rDNA over B rDNA after several vegetative generations. Therefore when taken in a physiological context, the differences in the affinity of ssA-TIBF for the C3 versus B extended Type I repeat sequences *in vitro* are sufficient to explain the replicative differences observed between the rDNAs *in vivo*. From these data, I propose that ssA-TIBF could be involved in the regulation of rDNA replication.

Sequence-specific single-stranded DNA binding factors have increasingly been reported to have roles in transcriptional regulation and recombination (14, 26, 39, 40, 59, 64, 84, 85, 105) in both lower and higher eukaryotes. Single-stranded DNA binding factors that recognize sequences within chromosomal replication origin regions in eukaryotes have also been identified. For example the yeast ARS consensus sequence binding protein (ACBP) specifically recognizes the T-rich strand of the 11 nucleotide ARS core (54, 65, 86). ACBP is sensitive to mutations that eliminate origin function in *in vitro* replication systems. Another mammalian single-stranded DNA binding protein, Pur, has been described which binds specifically to conserved purine rich sequence elements found at several origins of replication and certain gene flanking region (9). Pur has been shown to specifically bind elements in two zones of DNA replication initiation of the genes *c-myc* and *dhfr*.

ssA-TIBF also specifically recognizes the Type I repeat that is an essential component of the promoter *in vitro*, despite sequence differences at 8 nucleotide positions within the repeat and different 3' flanking sequences. ssA-TIBF recognizes the extended form of the Type I repeat promoter element with an affinity equal to that for the extended upstream Type I repeats. This suggests that ssA-TIBF may function in both rDNA replication and transcription. Taken together, our data suggests that ssA-TIBF may be a transcription factor that is involved the regulation of rDNA replication.

Evidence coupling DNA replication and transcription has been accruing over the last few years. This link may involve higher order chromatin structure. For instance, the transcriptional active regions of the genome tend to be replicated early in S phase (6, 50, 91), whereas transcriptionally inactive regions such as telomeres (74) and the inactivated X chromosome (43) generally are replicated late in S phase. One possible explanation for the coupling of DNA replication and transcription is to coordinate the direct competition between the nucleosomes and factors involved in DNA replication and transcription.

Specific coupling of transcription factor binding sites and DNA replication has also been observed (7). The replication of many mammalian viral genomes requires or is stimulated by transcription factor binding sites (reviewed in (30)). In the case of SV40 DNA replication, the presence of auxillary sequences containing binding sites for several transcription factors enhances replication up to 100 fold (29, 48, 53, 68, 70). However, the auxillary sequences must be in

close proximity to the core origin to exert their influence (21, 57, 68). Similar template requirements have been shown to be necessary for polyomavirus DNA replication (49). One possible explanation for the necessarily tight association of the auxillary sequences and the core origin is that direct protein-protein interactions between the transcription factors and the replicative machinery are responsible for the increase in SV40 DNA replication. Direct protein-protein interactions with transcription factors may either stabilize or increase the activity of the replication machinery.

In some cases, stimulation of DNA replication by transcription factors has been shown to require the activation domains of the transcription factor. In both polyomavirus and SV40 DNA replication, when wildtype transcription factor binding sites were replaced with binding sites for other transcription factors, only transcription factors with specific activation domains were able to stimulate DNA replication and inhibit repression during chromatin assembly (8, 24, 47). These data suggest that the activation of DNA replication and transcription may involve similar mechanisms.

In yeast the transcription factor ARS binding factor 1 (ABF1) has been shown to be important for ARS function. ABF1 binding sites occur at only a subset of origins, but the function of these origins in plasmid maintenance is reduced when ABF1 binding sites are interrupted (15, 33, 35, 88, 101). The role of ABF1 at the chromosome origin *ARS1* has been characterized (73). Binding of ABF1 is required for *ARS1* function. However, a related transcription factor that also interacts with

certain ARS elements, RAP1, is able to substitute for ABF1 when the ABF1 binding site is replaced with the RAP1 binding site. When the binding site for ABF1 is replaced with a binding site for a protein that lacks the activation domain, the functional contribution of the ABF1 site is lost.

Binding sites for transcription factors are associated with potential origins of replication in mammals. Transcriptional enhancer sequences and an origin of replication have been shown to colocalize upstream of the human *c-myc* gene (56). Transcription factor binding sites are also located adjacent to the origin of bi-directional replication of the *dhfr* locus of Chinese Hamster Ovary Cells (reviewed in 52). These binding sites include Oct1, and AP1.

ssA-TIBF may be related to UBF, an essential RNA polymerase I transcription factor that specifically interacts with the rRNA upstream promoter elements (UPE) of *Xenopus* (3) mouse (2), rat (89) and human (4). A complex between hUBF and another factor, hSL1 is required to activate RNA polymerase I transcription (4). Two related forms of UBF, UBF1 and UBF2, exhibit the molecular mass of 97 and 94 kD, respectively, in mouse, rat and human. This is very similar to the molecular mass of ssA-TIBF of ~95 kD.

Evidence suggests that UBF binds to the UPE as a dimer, or is capable of dimerizing after binding (79). The possibility exists that ssA-TIBF may form a dimer since it elutes from a gel filtration column with a molecular mass of 250 kD. However, a molecular mass of 250 kD would suggest a heteromultimer for ssA-TIBF.

UBF is a phosphoprotein and treatment with phosphatase reduces its ability to activate transcription *in vitro* and alters its cellular distribution (79). I have speculated that the formation of multiple shifted bands corresponding to ssA-TIBF-DNA complexes in gel mobility shift assays with crude cellular extracts may be due to degradation products or modifications of ssA-TIBF (see Paper 1). One possibility is that ssA-TIBF is a phosphoprotein, therefore it would be interesting to test the phosphorylation state of ssA-TIBF.

UBF contains four domains, HMG boxes, with homology to high mobility group proteins 1 and 2 (HMG1 and HMG2) (60). UBF2 differs from UBF1 by a 37 amino acid deletion in the HMG box 2. A truncated version of UBF2, lacking N-terminal sequences, is deficient in UPE binding and dimerization while truncated UBF1 is not. This suggests that the HMG box 2 may be involved in these functions (79).

HMGs are abundant nonhistone chromosomal proteins found in many eukaryotes (77). HMGs are also generally small (~20 kD) nonspecific DNA-binding proteins. HMG1 and 2 belong to a subset of HMGs and are thought to be involved in dynamic changes in chromatin structure during the cell cycle. A *Tetrahymena* protein, HMG B, also shows homology with HMG1 and 2 and is associated with both the macronucleus and the micronucleus (103). HMG B mRNA is specifically induced early in conjugation. During this period the micronucleus undergoes a series of changes requiring DNA replication, recombination, repair, and transcription. One possible conclusion is that HMG B is involved in the chromatin structural changes necessary for a broad

range of micronuclear and macronuclear processes, including DNA replication and transcription.

In addition to UBF, other proteins that contain HMG boxes, conserved DNA binding elements with sequence similarity to HMG 1 and 2, have been identified (77). Some of these factors with HMG boxes are sequence-specific transcription factors. This suggests that HMG boxes may be a common DNA-binding motif.

Model: Roles of ssA-TIBF, ssT-TIBF, and ds-TIBF in rDNA replication

Sequences that have been shown to be involved in rDNA replication, the Type I repeats, are distributed throughout an ~600 bp region in the 5' NTS (67). This structure collaborates with a proposed model suggesting a modular control in eukaryotic DNA replication in which several, simple, repetitive DNA elements are responsible for origin function when they are appropriately arrayed (7, 41). These modular elements include DNA unwinding elements, pyrimidine tracts, scaffold-associated regions, transcription factor binding sites, ARS-like sequences, and alternative DNA structures, such as bent DNA. Several of these DNA elements have been found within the 5' NTS of the rDNA molecule (D. Larson, unpublished data). Thus, it is possible that the control of rDNA replication involves these modular element.

ssT-TIBF is a nonspecific single-stranded DNA-binding that binds to T-rich sequences in general. As noted previously ssT-TIBF resembles RP-A, a nonspecific, single-stranded DNA binding protein which has a

preference for pyrimidine-rich sequences (62). RP-A is essential for SV40 DNA replication *in vitro*. Thus the role of ssT-TIBF in DNA replication may be similar to that of RP-A, stabilizing unwound DNA and stimulating the DNA replication machinery through protein-protein interactions.

ds-TIBF binds to duplexed Type I repeat in *in vitro* assays with modest specificity but with high affinity; ds-TIBF prefers to interact with AT-rich sequences. These results do not address ds-TIBFs binding specificity *in vivo*. Nor do these results exclude the possibility that ds-TIBF may play a physiological role *in vivo*. Since ds-TIBF recognizes only duplexed DNA sequences, this suggests that ds-TIBF may play a negative regulatory role in rDNA replication. For DNA replication to proceed, single-stranded sequences must be available. The binding of ds-TIBF to the Type I repeat may inhibit the unwinding event normally induced by the 5' NTS modular elements, thus regulating rDNA replication by the availability of the single-stranded initiation template. In this eventuality rDNA replication would require that ds-TIBF be displaced prior to the unwinding process. A negatively regulatory role of ds-TIBF would suggest that ds-TIBF may be under cell cycle control. Such a scenario would require that ds-TIBF inhibition of rDNA replication be inactive during cellular S phase.

ssA-TIBF is a single-stranded DNA-binding protein that specifically interacts with the A-rich strand of the Type I repeat. The function of ssA-TIBF in rDNA replication may involve the specific recognition of unwound sequences within the 5' NTS. The sequence

specific recognition of Type I sequences by ssA-TIBF may be important for directing the replication machinery to the 5' NTS near the center of the rDNA palindrome, possibly enabling the rapid replication of the entire molecule from a single origin. ssA-TIBF may then function to recruit the replication machinery to the origin region. The role of ssA-TIBF may also involve direct stimulation of DNA replication through protein-protein contacts with the replication machinery. The preference of ssA-TIBF for C3 vs B rDNA sequences in gel mobility shift assays mimics the replication ability of these rDNAs *in vivo*. These results corroborates with a positive regulatory role of ssA-TIBF in rDNA replication.

ssA-TIBF also interacts with the Type I repeat that is an essential element of the promoter *in vitro* (75). This suggests that ssA-TIBF may also be an RNA polymerase I transcription factor capable of stimulating rDNA transcription as well as replication. Experiments that may elucidate the exact role of ssA-TIBF are described in the following section.

Future direction

Further experiments are necessary to more fully understand the significance of the differential recognition of C3 versus B Domain 2 Type I repeats by ssA-TIBF. One way to accomplish this is to compare the DNase I footprinting patterns that result when the two different Type I repeat plus 3' flanking sequences are used as binding substrates. It will also be important to compare the promoter element Type I

repeat with the upstream Type I repeat sequences in the context of their 3' flanking sequences. Hydroxyl radical footprinting can more precisely map protein-DNA contact sites due to the smaller size of the reagent. Therefore, hydroxyl radical footprinting can be used in addition to DNase I footprinting to develop a more detailed understanding of the specific nucleotides contacted by ssA-TIBF.

Other experiments are necessary to determine the functional role of ssA-TIBF in rDNA replication. This characterization will be facilitated by cloning ssA-TIBF. A partial amino acid sequence of ssA-TIBF can be obtained from purified ssA-TIBF and used to screen a *Tetrahymena thermophila* cDNA library (provided by Dr. Tohru Takemasa). Comparison of the cloned sequence of ssA-TIBF with other proteins in the database may reveal a homology that could suggest a functional role for ssA-TIBF. It would be interesting to look for HMG boxes in the amino acid sequence of ssA-TIBF due to the potential relationship of ssA-TIBF to UBF.

Another approach to ascertain the function of ssA-TIBF in rDNA replication would be to test the role of ssA-TIBF in an *in vitro* DNA replication system. Antibodies raised against purified ssA-TIBF could be used to inhibit ssA-TIBF in *Tetrahymena* cellular extracts used in an *in vitro* DNA replication system. This may determine whether ssA-TIBF is essential for rDNA replication *in vitro*. Eventually, an *in vitro* system of DNA replication can be used to determine other factors involved in rDNA replication and to study the protein-protein interactions that are necessary for rDNA replication.

Similarly, it would be interesting to test the role of ssA-TIBF in an *in vitro* transcription system; to determine the functional significance of the binding of ssA-TIBF to the Type I repeat that is an essential promoter element *in vitro* (75).

Once the function of ssA-TIBF has been determined *in vitro*, it would be necessary to correlate these results with *in vivo* studies. Depending on the copy number of the gene for ssA-TIBF in the macronucleus, It may be possible to transform the *Tetrahymena* macronucleus with a mutant ssA-TIBF gene by homologous recombination. This will be technically difficult, requiring the microinjection of the gene into the developing anlagen, the macronuclear precursor formed after mating. The pRD4-1 vector, previously used to transform *Tetrahymena* with the telomerase RNA gene (117), can be used in these studies. pRD4-1 carries cis-acting replication control elements and a paromomycin resistant marker in a complete copy of the rDNA. Other transformation vectors that may also be used are being constructed in Drena Larson's and Eric Henderson's labs.

Finally, depending on the turnover rate of the ssA-TIBF protein in the macronucleus, anti-sense DNA could be microinjected into the macronucleus. Another approach would be to transform the macronucleus with an inverted copy of the ssA-TIBF gene in an attempt to constitutively generate antisense RNA. If ssA-TIBF is essential, the microinjection of anti-sense RNA should lead to a decreased growth rate and perhaps cell death. These experiments in addition to the

experiments listed above should elucidate the role of ssA-TIBF in *Tetrahymena thermophila* rDNA replication.

LITERATURE CITED

1. **Altman, S.** 1990. Ribonuclease P. J. Biol. Chem. **265**:20053-20056.
2. **Bell, S., H.-M. Jantzen, and R. Tjian.** 1990. Assembly of alternative multiprotein complexes directs rRNA promoter selectivity. Genes & Dev. **4**:943-954.
3. **Bell, S., C. Pikaard, R. Reeder, and R. Tjian.** 1989. Molecular mechanisms governing species-specific transcription of ribosomal RNA. Cell **59**:489-497.
4. **Bell, S.P., R.M. Learned, H.-M. Jantzen, and R. Tjian.** 1988. Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. Science **241**:1192-1197.
5. **Bell, S.P. and B. Stillman.** 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature **357**:128-134.
6. **Benard, M. and G. Pierron.** 1992. Mapping of a *Physarum* chromosomal origin of replication tightly linked to developmentally-regulated profilin gene. Nucleic Acids Research **20**:3309-3315.

7. **Benbow, R.M., J. Zhao, and D.D. Larson.** 1992. On the nature of origins of DNA replication in eukaryotes. *BioEssays* **10**:661-670.

8. **Bennett-Cook, E.R. and J.A. Hassell.** 1991. Activation of polyomavirus DNA replication by yeast GAL4 is dependent on its transcriptional activation domains. *EMBO J.* **10**:959-969.

9. **Bergemann, A.D. and E.M. Johnson.** 1992. The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. *Mol. Cell. Biol.* **12**:1257-1265.

10. **Bonven, B.J., E. Gocke, and O. Westergaard.** 1985. A high affinity topoisomerase I binding sequence is clustered at DNase I hypersensitive sites in *Tetrahymena* r-chromatin. *Cell* **41**:541-551.

11. **Brewer, B.J. and W.L. Fangman.** 1988. A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* **55**:637-643.

12. **Brill, S.J. and B. Stillman.** 1989. Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. *Nature* **324**:92-95.

13. **Brown, G.W., T.E. Melendy, and D.S. Ray.** 1992. Conservation of structure and function of DNA replication protein A in the trypanosomatid *Crithidia fasciculata*. *Proc. Natl. Acad. Sci. USA* **89**:10227-10231.
14. **Brunel, F., P.M. Alzari, P. Ferrara, and M.M. Zakin.** 1991. Cloning and sequencing of PYBP, a pyrimidine-rich specific single strand DNA-binding protein. *Nucleic Acids Res.* **19**:5237-5245.
15. **Buchman, A.R. and R.D. Kornberg.** 1990. A yeast *ARS*-binding protein activates transcription synergistically in combination with other weak activating factors. *Mol. Cell. Biol.* **10**:887-897.
16. **Campbell, J.L. and C.S. Newlon.** 1991. Chromosomal DNA replication, p. 41-146. *In* J.R. Broach, J.R. Pringle, and E.W. Jones (ed.), *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis and Energetics*. Cold Spring Harbor Press, Cold Spring Harbor, New York.
17. **Carminati, J.L., C.G. Johnston, and T.L. Orr-Weaver.** 1992. The *Drosophila* ACE3 chorion element autonomously induces amplification. *Mol. Cell. Biol.* **12**:2444-2453.

18. **Cech, T.R.** 1986. Ribosomal RNA gene expression in *Tetrahymena*: Transcription and RNA splicing, p. 203-225. In J.G. Gall (ed.), The molecular biology of ciliated protozoa. Academic Press, Inc., Orlando.

19. **Cech, T.R. and S.L. Brehm.** 1981. Replication of the extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila*. Nucleic Acids Res. **9**:3531-3543.

20. **Challoner, P.B., A.A. Amin, R.E. Pearlman, and E.H. Blackburn.** 1985. Conserved arrangements of repeated DNA sequences in nontranscribed spacers of ciliate ribosomal RNA genes: evidence for molecular coevolution. Nucleic Acids Res. **13**:2661-2680.

21. **Chandrasekharappa, S.C. and K.N. Subramanian.** 1987. Effects of positions and orientation of the 72-base-pair-repeat transcriptional enhancer on replication from the simian virus 40 core origin. J. Virol. **61**:2973-2980.

22. **Chang, D.D. and D.A. Clayton.** 1987. A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. Science **235**:1178-1184.

23. **Chang, D.D. and D.A. Clayton.** 1987. A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J.* **6**:409-417.
24. **Cheng, L., J.L. Workman, R.E. Kingston, and T.J. Kelly.** 1992. Regulation of DNA replication *in vitro* by the transcriptional activation domain of GAL4-VP16. *Proc. Natl. Acad. Sci. USA* **89**:589-593.
25. **Clayton, D.A.** 1991. Replication and transcription of vertebrate mitochondrial DNA. *Annu. Rev. Cell Biol.* **7**:453-478.
26. **Collick, A., M.G. Dunn, and A.J. Jeffreys.** 1991. Minisatellite binding protein Msp-1 is a sequence-specific single-stranded DNA-binding protein. *Nucleic Acids Res.* **19**:6399-6404.
27. **Dalley, L., M.S. Caddle, N. Heintz, and N.H. Heintz.** 1990. Purification of RIP60 and RIP100, mammalian proteins with origin-specific DNA-binding and ATP-dependent DNA helicase activities. *Mol. Cell. Biol.* **10**:6225-6235.
28. **Davis, T.L., A.B. Firulli, and A.J. Kinniburgh.** 1989. Ribonucleoprotein and protein factors bind to an H-DNA-forming *c-myc* DNA element: Possible regulators of the *c-myc* gene. *Proc. Natl. Acad. Sci. USA* **86**:9682-9686.

29. **DeLucia, A.L., S. Deb, K. Partin, and P. Tegtmeyer.** 1986. Functional interactions of the simian virus 40 core origin of replication with flanking regulatory sequences. *J. Virol.* **57**:138-144.

30. **DePamphilis, M.L.** 1988. Transcriptional elements as components of eukaryotic origins of DNA replication. *Cell* **52**:635-638.

31. **Deshpande, A.M. and C.S. Newlon.** 1992. The *ARS* consensus sequence is required for chromosomal origin function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**:4305-4313.

32. **Diffley, J.F.X. and J.H. Cocker.** 1992. Protein-DNA interactions at a yeast replication origin. *Nature* **357**:169-172.

33. **Diffley, J.F.X. and B. Stillman.** 1988. Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc. Natl. Acad. Sci. USA* **85**:2120-2124.

34. **Dubey, D.D., L.R. Davis, S.A. Greenfeder, L.Y. Ong, J. Zhu, J.R. Broach, C.S. Newlon, and J.A. Huberman.** 1991. Evidence suggesting that the *ARS* elements associated with silencers of the yeast mating type locus, *HML*, do not function as chromosomal DNA replication origins. *Mol. Cell. Biol.* **11**:5346-5355.

35. **Eisenberg, S., C. Civalier, and B.K. Tye.** 1988. Specific interaction between a *Saccharomyces cerevisiae* protein and a DNA element associated with certain autonomously replicating sequences. *Proc. Natl. Acad. Sci. USA* **85**:743-746.

36. **Fairman, M.P. and B. Stillman.** 1988. Cellular factors required for multiple stages of SV40 replication *in vitro*. *EMBO J.* **7**:1211-1218.

37. **Fangman, W.L. and B.J. Brewer.** 1991. Activation of replication origins within yeast chromosomes. *Annu. Rev. Cell Biol.* **7**:375-402.

38. **Ferguson, B.M. and W.L. Fangman.** 1992. A position effect on the time of replication origin activation in yeast. *Cell* **68**:333-339.

39. **Finocchiaro, L.M.E., P. Amati, and G.C. Glikin.** 1991. Single strand binding protein specific for the polyoma early-coding strand of PEA1 (AP1) regulatory sequence. *Nucleic Acids Res.* **19**:4279-4287.

40. **Gaillard, C. and F. Strauss.** 1990. Sequence-specific single-strand-binding protein for the Simian Virus 40 early promoter stimulates transcription *in vitro*. *J. Mol. Biol.* **215**:245-255.

41. **Gale, J.M., R.A. Tobey, and J.A. D'Anna.** 1992. Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells. *J. Mol. Biol.* **224**:343-358.

42. **Galli, I., S.M.M. Iguchi-Arigo, and H. Arigo.** 1992. The AT-rich tract of the SV40 ori core: negative synergism and specific recognition by single stranded and duplex DNA binding proteins. *Nucleic Acids Res.* **20**:3333-3339.

43. **Grant, S.G. and V.M. Chapman.** 1988. Mechanisms of X-chromosome regulation. *Annu. Rev. Genet.* **22**:199-233.

44. **Greenfeder, S.A. and C.S. Newlon.** 1992. A replication map of a 61-kb circular derivative of *Saccharomyces cerevisiae* chromosome III. *Mol. Cell. Biol.* **13**:999-1013.

45. **Greider, C.W. and E.H. Blackburn.** 1987. The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51**:887-898.

46. **Greider, C.W. and E.H. Blackburn.** 1989. A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* **337**:331-337.

47. **Guo, Z.-S. and M.L. DePamphilis.** 1992. Specific transcription factors stimulate simian virus 40 and polyomavirus origins of DNA replication. *Mol. Cell. Biol.* **12**:2514-2524.
48. **Guo, Z.-S., C. Gutierrez, U. Heine, J.M. Sogo, and M.L. DePamphilis.** 1989. Origin auxiliary sequences can facilitate initiation of simian virus 40 DNA replication *in vitro* as they do *in vivo*. *Mol. Cell. Biol.* **9**:3593-3602.
49. **Hassell, J.A., W.J. Muller, and C.R. Muelier.** 1986. The dual role of the polyomavirus enhancer in transcription and DNA replication. *Cancer Cells (Cold Spring Harbor)* **4**:561-569.
50. **Hatton, K.S., V. Dhar, E.H. Brown, M.A. Iqbal, S. Stuart, V.T. Didamo, and C.L. Schildkraut.** 1988. Replication program of active and inactive multigene families in mammalian cells. *Molecular and Cellular Biology* **8**:2149-2158.
51. **Heck, M.M.S. and A.C. Spradling.** 1990. Multiple replication origins are used during *Drosophila* chorion gene amplification. *J. Cell Biol.* **110**:903-914.
52. **Held, P.G. and N.H. Heintz.** 1992. Eukaryotic replication origins. *Biophys. Acta.* **1130**:235-246.

53. **Hertz, G.Z. and J. Mertz.** 1986. Bidirectional promoter elements of simian virus 40 are required for efficient replication of the viral DNA. *Mol. Cell. Biol.* **6**:3513-3522.
54. **Hofmann, J.F.-X. and S.M. Gasser.** 1991. Identification and purification of a protein that binds the yeast *ARS* consensus sequence. *Cell* **64**:951-960.
55. **Huberman, J.A., J. Zhu, L.R. Davis, and C.S. Newlon.** 1988. Close association of a DNA replication origin and an *ARS* element on chromosome III of the yeast, *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **16**:6373-6383.
56. **Iguchi-Ariga, S.M.M., T. Okazaki, T. Itani, M. Ogata, Y. Sato, and H. Ariga.** 1988. An initiation site of DNA replication with transcriptional enhancer activity present upstream of the *c-myc* gene. *The EMBO Journal* **7**:3135-3142.
57. **Innis, J.W. and W.A. Scott.** 1984. DNA replication and chromatin structure of simian virus insertion mutants. *Mol. Cell. Biol.* **4**:1499-1507.
58. **Ishimi, Y., A. Claude, P. Bullock, and J. Hurwitz.** 1988. Complete enzymatic synthesis of DNA containing the SV40 origin of replication. *J. Biol. Chem.* **263**:19723-19733.

59. **Jansen-Durr, P., M. Boshart, B. Lupp, A. Bosserhoff, R.W. Frank, and G. Schutz.** 1992. The rat poly pyrimidine tract binding protein (PTB) interacts with a single-stranded DNA motif in a liver-specific enhancer. *Nucleic Acids Res.* **20**:1243-1249.
60. **Jantzen, H.-M., A. Admon, S.P. Bell, and R. Tjian.** 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **344**:830-836.
61. **Karwan, R., J.L. Bennett, and D.A. Clayton.** 1991. Nuclear RNase MRP processes RNA at multiple discrete sites: interaction with an upstream G box is required for subsequent downstream cleavages. *Genes. Dev.* **5**:1264-1276.
62. **Kim, C., R.O. Snyder, and M.S. Wold.** 1992. Binding properties of replication protein A from human and yeast cells. *Mol. Cell. Biol.* **12**:3050-3059.
63. **Kiss, T. and W. Filipowicz.** 1992. Evidence against a mitochondrial location of the 7-2/MRP RNA in mammalian cells. *Cell* **70**:11-16.
64. **Kolluri, R., T.A. Torrey, and A.J. Kinniburgh.** 1992. A CT promoter element binding protein: definition of a double-strand and a novel single-strand DNA binding motif. *Nucleic Acids Res.* **20**:111-116.

65. **Kuno, K., S. Murakami, and S. Kuno.** 1990. Single-strand-binding factor(s) which interact with *ARS1* of *Saccharomyces cerevisiae*. *GENE* **95**:73-77.
66. **Lamond, A.I., S.M.L. Barabino, and B.J. Blencowe.** 1990. The mammalian pre-mRNA splicing apparatus, p. 243-257. *In* E. F. and D.M.J. Lilley (ed.), *Nucleic Acids and Molecular Biology*. Springer-Verlag, Berlin.
67. **Larson, D.D., E.H. Blackburn, P.C. Yaeger, and E. Orias.** 1986. Control of rDNA replication in *Tetrahymena* involves a *cis*-acting upstream repeat of a promoter element. *Cell* **47**:229-240.
68. **Lee-Chen, G.-J. and M. Woodworth-Gutai.** 1986. Simian virus 40 DNA replication: Functional organization of regulatory elements. *Mol. Cell. Biol.* **6**:3086-3093.
69. **Li, J. and T.J. Kelly.** 1985. Simian virus 40 DNA replication in vitro: Specificity of initiation and evidence for bidirectional replication. *Mol. Cell. Biol.* **5**:1238-1246.
70. **Li, J.J., K.W.C. Peden, R.A.F. Dixon, and T. Kelly.** 1986. Functional organization of the simian virus 40 origin of DNA replication. *Mol. Cell. Biol.* **6**:1117-1128.

71. **Linskens, M.H.K. and J.A. Huberman.** 1988. Organization of replication of ribosomal DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:4927-4935.

72. **Malkas, L.H. and E.F. Baril.** 1989. Sequence recognition protein for the 17-base-pair A+T-rich tract in the replication origin of simian virus 40 DNA. Proc. Natl. Acad. Sci. USA **86**:70-74.

73. **Marahrens, Y. and B. Stillman.** 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science **255**:817-823.

74. **McCarroll, R.M. and W.L. Fangman.** 1988. Time of replication of yeast centromeres and telomeres. Cell **54**:505-513.

75. **Miyahara, K., N. Hashimoto, T. Higashinakagawa, and R.E. Pearlman.** in Press. Common sequence elements are important for transcription and replication of the extrachromosomal rRNA genes of *Tetrahymena*. Gene

76. **Morrison, A., H. Araki, A.B. Clark, R.K. Hamatake, and A. Sugino.** 1990. A third essential DNA polymerase in *S. cerevisiae*. Cell **62**:1143-1151.

77. **Ner, S.S.** 1992. HMGs everywhere: a new and diverse class of DNA-binding proteins is identified by a common sequence motif. *Curr. Biol.* **4**:208-210.
78. **Nishida, C., P. Reinhard, and S. Linn.** 1988. DNA repair synthesis in human fibroblasts requires DNA polymerase δ . *J. Biol. Chem.* **263**:501-510.
79. **O'Mahony, D.J., S.D. Smith, W. Xie, and L.I. Rothblum.** 1992. Analysis of the phosphorylation, DNA-binding and dimerization properties of the RNA polymerase I transcription factors UBF1 and UBF2. *Nucl. Acids Res.* **20**:1301-1308.
80. **Palen, T.E. and T.R. Cech.** 1984. Chromatin structure at the replication origins and transcription initiation regions of the ribosomal RNA genes of *Tetrahymena*. *Cell* **36**:933-942.
81. **Pan, W.C., E. Orias, M. Flacks, and E.H. Blackburn.** 1982. Allele-specific, selective amplification of a ribosomal RNA gene in *Tetrahymena thermophila*. *Cell* **28**:595-604.
82. **Prelich, G., M. Kostura, D.R. Marshak, M.B. Mathews, and B. Stillman.** 1987. The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication *in vitro*. *Nature* **326**:471-475.

83. **Rivier, D.H. and J. Rine.** 1992. An origin of DNA replication and a transcription silencer require a common element. *Science* **256**:659-663.

84. **Roller, R.J., A.L. McCormick, and B. Roizman.** 1989. Cellular proteins specifically bind single- and double-stranded DNA and RNA from the initiation site of a transcript that crosses the origin of DNA replication of herpes simplex virus 1. *Proc. Natl. Acad. Sci. USA* **86**:6518-6522.

85. **Santoro, I.M., T.-M. Yi, and K. Walsh.** 1991. Identification of single-stranded-DNA-binding proteins that interact with muscle gene elements. *Mol. Cell. Biol.* **11**:1944-1953.

86. **Schmidt, A.M.A., S.U. Herterich, and G. Krauss.** 1991. A single-stranded DNA binding protein from *S. cerevisiae* specifically recognizes the T-rich strand of the core sequence of ARS elements and discriminates against mutant sequences. *EMBO J.* **10**:981-985.

87. **Shinomiya, T. and S. Ina.** 1991. Analysis of chromosomal replicons in early embryos of *Drosophila melanogaster* by two-dimensional gel electrophoresis. *Nucleic Acids Res.* **19**:3935-3941.

88. **Shore, D., D.J. Stillman, A.H. Brand, and K.A. Nasmyth.** 1987. Identification of silencer binding proteins from yeast: possible roles in *SIR* control and DNA replication. *EMBO J.* **6**:461-467.

89. **Smith, S., E. Oriahi, D. Lowe, H.-F. Yang-Yen, D. O'Mahony, K. Rose, K. Chen, and L. Rothblum.** 1990. Characterization of factors that direct transcription of rat ribosomal DNA. *Mol. Cell. Biol.* **10**:3105-3116.

90. **So, A.G. and K.M. Downey.** 1992. Eukaryotic DNA replication. *Crit. rev. Biochem. Mol. Biol.* **27**:129-155.

91. **Spack, E.G., E.D. Lewis, B. Paradowski, R.T. Schimke, and P.P. Jones.** 1992. Temporal order of DNA replication in the H-2 major histocompatibility complex of the mouse. *Mol. Cell. Biol.* **12**:5174-5188.

92. **Stahl, H., P. Droge, and R. Knippers.** 1986. DNA helicase activity of SV40 large tumor antigen. *EMBO J.* **5**:1939-1944.

93. **Stillman, B.** 1989. Initiation of Eukaryotic DNA Replication *In Vitro*. *Annu. Rev. Cell Biol.* **5**:197-245.

94. **Stillman, B.W. and Y. Gluzman.** 1985. Replication and supercoiling of simian virus 40 DNA in cell extracts from human cells. *Mol. Cell. Biol.* **5**:2051-2060.

95. **Stohl, L.L. and D.A. Clayton.** 1992. *Saccharomyces cerevisiae* contains an RNase MRP that cleaves at a conserved mitochondrial RNA sequence implicated in replication priming. *Mol. Cell. Biol.* **12**:2561-2569.

96. **Tjian, R.** 1978. The binding site on SV40 DNA for a T antigen-related protein. *Cell* **13**:165-179.

97. **Topper, J.N., J.L. Bennett, and D.A. Clayton.** 1992. A role for RNAase MRP in mitochondrial RNA processing. *Cell* **70**:16-20.

98. **Tsurimoto, T. and B. Stillman.** 1989. Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, α and δ . *EMBO J.* **8**:3883-3889.

99. **Vassilev, L. and E.M. Johnson.** 1990. An initiation zone of chromosomal DNA replication located upstream of the *c-myc* gene in proliferating HeLa cells. *Mol. Cell. Biol.* **10**:4899-4904.

100. **Vaughn, J.P., P.A. Dijkwel, and J.L. Hamlin.** 1990. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. *Cell* **61**:1075-1087.
101. **Walker, S.S., S.C. Francesconi, and S. Eisenberg.** 1990. A DNA replication enhancer in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **87**:4665-4669.
102. **Walker, S.S., A.K. Malik, and S. Eisenberg.** 1991. Analysis of the interactions of functional domains of a nuclear origin of replication from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **19**:6255-6262.
103. **Wang, T. and C.D. Allis.** 1993. An abundant high-mobility-group-like protein is targeted to micronuclei in a cell cycle-dependent and developmentally regulated fashion in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **13**:163-173.
104. **Weinberg, D.H., K.L. Collins, P. Simancek, A. Russo, M.S. Wold, D.M. Virship, and T.J. Kelly.** 1990. Reconstitution of SV40 DNA replication with purified proteins. *Proc. Natl. Acad. Sci. USA* **87**:8692-8696.

105. **Wilkison, W.O., H.Y. Min, K.P. Claffey, B.L. Satterberg, and B.M. Spiegelman.** 1990. Control of the adipsin gene is adipocyte differentiation. *J. Biol. Chem.* **265**:477-482.
106. **Wobbe, C.R., F. Dean, L. Weissbach, and J. Hurwitz.** 1985. *In vitro* replication of duplex circular DNA containing the simian virus 40 DNA origin site. *Proc. Natl. Acad. Sci. USA* **82**:5710-5714.
107. **Wold, M.S. and T.J. Kelly.** 1988. Purification and characterization of replication protein A, a cellular protein required for *in vitro* replication of simian virus 40 DNA. *Proc. Natl. Acad. Sci. USA* **85**:2523-2527.
108. **Wold, M.S., J. Li, D.H. Weinberg, D.M. Vershup, E. Verkeyer, and T. Kelly.** 1988. Cellular proteins required for SV40 DNA replication *in vitro*. *Cancer Cells* **6**:133-144.
109. **Wold, M.S., J.J. Li, and T.J. Kelly.** 1987. Initiation of simian virus 40 DNA replication *in vitro*: Large-tumor-antigen and origin-dependent unwinding of the template. *Proc. Natl. Acad. Sci. USA* **84**:3643-3647.
110. **Wong, T.W. and D.A. Clayton.** 1986. DNA primase of human mitochondria is associated with structural RNA that is essential for enzymatic activity. *Cell* **45**:817-825.

111. **Yaeger, P.C., E. Orias, W.-L. Shaiu, D.D. Larson, and E.H. Blackburn.** 1989. The replication advantage of a free linear rRNA gene is restored by somatic recombination in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **9**:452-460.
112. **Yang, L., M.S. Wold, J.J. Li, T.J. Kelly, and L.F. Liu.** 1987. Roles of DNA topoisomerases in SV40 DNA replication *in vitro*. *Proc. Natl. Acad. Sci. USA* **84**:950-954.
113. **Yao, M.-C.** 1986. Amplification of ribosomal RNA genes, p. 179-201. *In* J.G. Gall (ed.), The molecular biology of ciliated protozoa. Academic Press, Inc., Orlando, Florida.
114. **Young, L.S., H.M. Dunstan, P.R. Witte, T.P. Smith, S. Ottonello, and K.U. Sprague.** 1991. A class III transcription factor composed of RNA. *Science* **252**:542-546.
115. **Yu, G.-L. and E.H. Blackburn.** 1989. Transformation of *Tetrahymena thermophila* with a mutated circular ribosomal DNA plasmid vector. *Proc. Natl. Acad. Sci. U S A* **86**:8487-91.
116. **Yu, G.L. and E.H. Blackburn.** 1990. Amplification of tandemly repeated origin control sequences confers a replication advantage on rDNA replicons in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **10**:2070-2080.

117. Yu, G.L., J.D. Bradley, L.D. Attardi, and E.H. Blackburn. 1990. *In vivo* alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs. *Nature* **344**:126-132.
118. Zhu, J., C. Brun, H. Kurooka, M. Yanagida, and J.A. Huberman. 1992. Identification and characterization of a complex chromosomal replication origin in *Schizosaccharomyces pombe*. *Chromosoma* **102**:S7-S16.
119. Zhu, J., C.S. Newlon, and J.A. Huberman. 1992. Localization of a DNA replication origin and termination zone on chromosome III of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**:4733-4741.

ACKNOWLEDGMENTS

I would like to thank Drena for her excellent guidance and counseling. I really appreciated Drena's patients, especially when I was a complete novice. I certainly learned a lot from Drena.

I would also like to thank Eric Henderson whose creative genius was always a bonus whenever I ran into technical difficulties. Eric adds his own unique element of craziness to the lab environment.

The editorial comments provided by Robert Benbow, Alan Myers, Phil Haydon, and Terese Cotton for the papers and my dissertation were greatly appreciated. I would also like to thank my POS committee members for being available for all sorts of technical research advice.

I am grateful for the research contributions of Wen-Ling Shaiu and Zita Sibenaller. Wen-Ling's hydroxyl radical footprinting experiment was an important aspect of the first paper and Zita's assistance during the purification of ssA-TIBF was exceptionally helpful.

Thanks also goes out to all of my lab buddies: Tom, Shawn, Ted, Zita, Vicki, Zhen, Joong-kook, Du, Hong, Meilin, Wen-Ling, Linda, Audra and of course I cannot forget Bruce. You partying dudes have provided an enjoyable learning experience both in and out of lab!!!

Much appreciation also goes out to Steve for helping me to relax when I needed it. Steve should also be commended for his invaluable assistance in completing the final printing of my dissertation. I couldn't have done it with out you, Honey!!!

Finally, the moral support and understanding provided by my parents gave me the courage to keep trying. Thanks Mom and Dad!!

APPENDIX: ds-TIBF MAY CONTAIN AN RNA COMPONENT

Preliminary results suggest that ds-TIBF is associated with an RNA which enhances the affinity of ds-TIBF for Type I repeat. Partially purified ds-TIBF has optimum binding to the Type I repeat in 400 mM NaCl as opposed to 1M in crude extracts (data not shown). This suggested that something present in crude extracts stabilizes the DNA-protein complex. To identify the components of ds-TIBF, extracts were treated with proteases and nucleases and tested for their ability to form the ds-TIBF-DNA complex by gel mobility shift assays. ds-TIBF is sensitive to two different proteases, Proteinase K and Pronase (data not shown). Extracts treated with these proteases could not form the shifted DNA-protein complex. Thus ds-TIBF has a protein component.

These experiments also demonstrated that ds-TIBF is sensitive to nuclease treatment. In gel shift assays, crude cellular extracts treated with micrococcal nuclease, an enzyme which cleaves both DNA and RNA, resulted in a reduction of ds-TIBF binding ability (Fig. 1). However, DNase I did not affect the binding (data not shown). Treatment with RNase A, an enzyme which cleaves RNA, produced an aberrant shifting pattern (Fig. 1). In this case, two new bands were formed in addition to the band normally produced by untreated extracts. Taken together, these results suggest that ds-TIBF could contain an RNA component. Cleavage of this RNA may result in a reduction in binding or a conformational change in ds-TIBF which alters the mobility of the ds-TIBF-DNA complex in gel shift assays.

FIG. 1. Nuclease treatments affect the binding of ds-TIBF to the Type I repeat.

Gel mobility shift analysis of ds-TIBF binding after treating crude cellular extracts with nucleases. Input indicates the ^{32}P -labeled oligonucleotide in the absence of extract. The no nuclease lane depicts the shifted band without nuclease treatment. MNase, ds-TIBF-DNA complexes formed in extracts treated with .15 and 1.5 units micrococcal nuclease; RNase A, ds-TIBF-DNA complexes formed in extracts treated with .1 and 1 $\mu\text{g/ml}$ ribonuclease A.

Methods: 50 μl aliquots of crude *Tetrahymena* cellular extracts were treated for 15 min at room temperature with various amounts of nucleases. Micrococcal nuclease (MNase) treated extracts were incubated with a total of .15 or 1.5 Units MNase plus 5 mM CaCl_2 . Ribonuclease A (RNase A) treated extracts were incubated with a final concentration of .1 or 1 $\mu\text{g/ml}$ RNase A. 5 μl (4 μg protein) of each reaction was incubated with 1 ng ds Type I₃₃ in gel mobility shift reactions (see paper 1, Materials and Methods) for 15 min on ice. The MNase reactions were stopped with EGTA, a Ca^{++} ion chelator, prior to being incubated with the input oligonucleotide. The reaction mixture was then electrophoresed on a 10% polyacrylamide gel.

It is possible that an RNA may facilitate ds-TIBF binding to the Type I repeat by forming an RNA-DNA complex with the Type I repeat. RNase H is an enzyme which cleaves the RNA in an RNA-DNA duplex. Treatment of the ds-TIBF-Type I repeat complex formed in partially purified extracts with RNase H resulted in a reduced shift (Fig. 2). Regardless of the amount of RNase H added, the shifted complex was never completely destroyed (Fig. 2, data not shown). This suggests that a specific copurifying RNA stabilizes interactions between ds-TIBF and the Type I repeat, but is not strictly required for binding.

There are many possible roles for an RNA associated with ds-TIBF. As suggested previously, the RNA component may simply stabilize the association of ds-TIBF with the Type I repeat. Sensitivity to RNase H suggests that the RNA forms a complex with the Type I repeat. This could occur by either forming a duplex with the unwound DNA template or by forming a triplex. The former is unlikely since ds-TIBF specifically binds the duplexed form of the Type I repeat. Also, the reduced shift in extracts pretreated with nucleases suggests that ds-TIBF associates with the RNA prior to binding the Type I repeat. Therefore, it is unlikely that ds-TIBF is recognizing an RNA-DNA hybrid. This suggests that the RNA may be stabilizing ds-TIBF binding by forming a triplex with the DNA template. This possibility can be addressed by testing the ability of purified ds-TIBF, lacking the RNA component, to bind to a Type I repeat triplex.

Ribonucleoproteins (RNPs) are involved in many cellular functions involving nucleic acids metabolism, such as telomere elongation (45),

FIG. 2. The ds-TIBF-Type I repeat complex is sensitive to RNase H.

Gel mobility shift analysis of ds-TIBF after treating the ds-TIBF-Type I repeat complex with RNase H. - and + denote untreated and treated samples, respectively. Lanes 1 and 2, no RNase H; lane 3, 2 units RNase H; lane 4, 4 units RNase H.

Methods: The ds-TIBF-Type I repeat complex was allowed to form in gel mobility shift reaction mixtures containing 5 μ l of a ds-TIBF phosphocellulose fraction and 1 ng ds Type I₃₃ for 15 min on ice (see paper 1, Materials and Methods). 2 or 4 units RNase H was then added and the mixture was incubated for another 15 min at room temperature. Controls without RNase H were also incubated an additional 15 min at room temperature. The reaction mixture was then electrophoresed on a 5% polyacrylamide gel.

tRNA (1) and mRNA (66) processing, and mitochondrial DNA synthesis (25). For example, the template sequence for telomere elongation is an intrinsic RNA component of the telomerase enzyme (46, 117).

Ribonuclease P is an RNP required for the endonucleolytic cleavage of precursor tRNAs to form the 5' tRNA termini (1). Finally, RNPs are also involved mRNA splicing (66). RNPs may also be involved in the regulation of transcription. A factor with the properties of a ribonucleoprotein associates with a potential transcriptional regulatory element of the *c-myc* gene (28).

Mitochondrial DNA replicates by a displacement loop process in which replication initiates on one strand, the heavy-strand, proceeds through 2/3 of the genomic sequence before replication begins on the light-strand template. A single-stranded initiation site is required for light-strand replication to begin. The primer for the initiation of mitochondrial heavy-strand DNA synthesis is generated by the cleavage of RNA transcripts at a few specific sites. A site specific endoribonuclease was isolated in both mouse and human mitochondria which may be involved in the primer formation. This factor is a ribonucleoprotein, RNase MRP (for mitochondrial RNA processing) and the RNA component is essential (23, 25). A yeast RNase MRP homolog cleaves precisely at the only RNA to DNA transition site in the yeast mtDNA replication origin (95). However, the role of RNase MRP in mitochondria DNA replication is not yet resolved. It is important to note that, during mitochondrial RNase MRP purification, less than 1% of RNase MRP activity (61) and RNase MRP RNA (22) is present in

mitochondrial fractions and that most of the RNA component is found in the nucleolus ((63, 97) and references therein). Also, RNAs isolated from two plant species that are analogous to the RNA component of mammalian RNase MRP are enriched in the nucleoli (63). This suggests that RNase MRP may also function in the nucleoli, perhaps in ribosomal RNA processing.

In addition to RNPs, free RNAs that associate with protein complexes may also be involved in nucleic acid metabolism. For example, in humans, the ribosomal 5.8S RNA is associated with the DNA primase involved in the initiation of mitochondrial light-strand DNA replication (110). Finally, the silkworm *Bombyx mori* TFIIR is an RNA polymerase III transcription factor composed of RNA that is required for *in vitro* transcription (114). Three other transcription factors are also required for RNA polymerase III *in vitro* transcription. Thus, one possible role of TFIIR could be to stabilize the pre-initiation transcription complex.