Nucleic Acids Research, 1998, Vol. 26, No. 5 **1309–1316**

Precise mapping and characterization of the RNA primers of DNA replication for a yeast hypersuppressive petite by *in vitro* capping with guanylyltransferase

Tina Graves[+], Michael Dante[+], Lynn Eisenhour and Thomas W. Christianson*

Department of Microbiology, Southern Illinois University, Carbondale, IL 62901, USA

Received October 16, 1997; Revised and Accepted January 12, 1998

ABSTRACT

The active origins of DNA replication for yeast (*Saccharomyces cerevisiae*) mitochondrial DNA share 280 conserved base pairs and have a promoter. Since intact replication intermediates retain their initiating ribonucleotide triphosphate, we used guanylyltransferase to *in vitro* cap the replication intermediates present in restriction enzyme-cut DNA from an ori-5 hypersuppressive petite. Restriction mapping and RNA sequencing of these labeled intermediates showed that each DNA strand is primed at a single discrete nucleotide, that one primer starts at the promoter and that the other primer starts 34 nt away, outside the conserved region. Deoxyribonuclease digestion of the capped fragments left resistant RNA primers, which enabled identification of zones of transition from RNA to DNA synthesis. Some of the results contradict the prevailing model for priming at the yeast mitochondrial origins.

INTRODUCTION

The mitochondrial genome of baker's yeast *Saccharomyces cerevisiae* is ~80 kb long. This genome can suffer large deletions resulting in respiratory-deficient yeast, which are called petite (or rho$^-$) mutants. The retained DNA in most rho$^-$ petites is tandemly repeated and amplified (1). While most petites are recessive when crossed with grande (rho$^+$ or wild-type) yeast, suppressive petites produce a high percentage of petites in such crosses and hypersuppressive (or supersuppressive) petites produce mostly petite progeny. Hypersuppressive petites were shown to have small genomes containing one of three or four (depending on the parental strain) homologous regions called rep or active ori sequences (2,3). These ori/rep sequences are believed to be the normal origins for yeast mtDNA replication. This conclusion is based on the assumption that hypersuppressitivity is the result of the tandemly repeated ori/rep sequences successfully competing for the DNA replication machinery (1), although recent work suggests preferential segregation could be the basis of hypersuppressitivity (4,5).

The various yeast ori/rep sequences share 280 bp of homologous DNA (6). The ori/rep sequence is AT-rich but contains three GC-rich clusters, designated A, B and C (Fig. 1). A yeast mitochondrial promoter is immediately adjacent to the C cluster; RNA is transcribed from this promoter (7,8). In various species of mammals there are three conserved sequence blocks (CSB) downstream of the mitochondrial light strand promoter (9); CSB 2 has strong homology with the yeast C cluster and CSB 1 also has a homolog in the yeast ori/rep sequence (6).

A model for initiating yeast mitochondrial DNA (mtDNA) replication at an ori/rep sequence (10,11) has been developed on biochemical evidence and by analogy with replication of mammalian mtDNA (11–16). In mammalian mitochondria RNA transcribed from the light strand promoter primes D-loop DNA synthesis and presumably primes synthesis of the first strand of DNA during replication (13). The model of yeast replication maintains that RNA transcribed from the yeast ori/rep promoter similarly serves to prime the first (promoter) strand of DNA replication. Supporting this step of the model is the fact that there are several inactive (not hypersuppressive) ori/rep sequences which are homologous to active ori/rep sequences except for the presence of a short DNA insertion which disrupts the promoter (7,17). In mammalian mitochondria the various transition sites from primer RNA to DNA synthesis occur downstream of the promoter in the region of the CSBs (18). An enzyme called MRP has been isolated from mammalian mitochondria (16) which cleaves RNA *in vitro* near or at a number of these *in vivo* transition sites (16,19,20); one of these sites is adjacent to CSB 2. MRP is postulated to produce RNA primer 3′-OH ends *in vivo*. The model for yeast replication initiation maintains that a similar cleavage occurs on the transcript from the ori/rep promoter to generate the primer for promoter strand replication. When the S1 nuclease technique was used with several ori/rep petites (21) to identify RNA primer and DNA ends (Fig. 1) a RNA 5′-end was mapped to the promoter and the RNA–DNA transition was placed at the boundary of the C cluster. In support of the model a yeast MRP enzyme has been found that cleaves RNA *in vitro* at one of several nucleotides immediately upstream of the CSB 2 homolog, cluster

*To whom correspondence should be addressed. Tel: +1 618 453 3003; Fax: +1 618 453 8036; Email: christianson@micro.siu.edu

[+]Present address: Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, St Louis, MO, USA

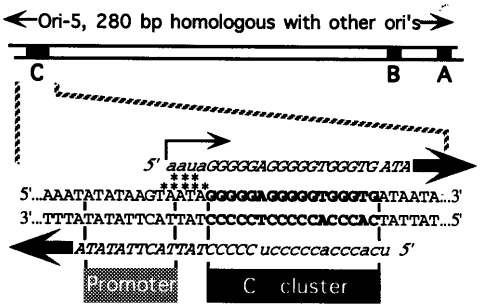

Figure 1. Map of ori-5 and the DNA sequence near its C cluster. The ori-5 priming sites and RNA–DNA transitions, as identified by Baldacci *et al.* (21), are shown. The template DNA sequence is in regular type, italics indicate newly replicated strands and the RNA primers attached to the replicated DNA are in lower case. Promoter strand primer synthesis starts at the promoter (bent arrow) and the transition to DNA synthesis is just 4 nt downstream. Non-promoter strand primer synthesis starts to the right of the C cluster and the transition to DNA synthesis is inside the C cluster. Yeast MRP cleavage sites (22) are designated by asterisks; two asterisks mark the preferred cleavage sites. Note that the leftmost MRP cleavage site could not function *in vivo* at ori-5, because it precedes the initiation site of the promoter.

C (22). As shown in Figure 1, the preferred cleavage site is 1 nt away from the RNA–DNA transition mapped by S1.

In mammalian mitochondria second strand RNA synthesis initiates after first strand replication has proceeded two thirds of the way around the mtDNA. The displaced single strand is capable of forming a stem and loop which serves as the site for priming second strand synthesis (14). The second strand RNA primer is synthesized by a mitochondrial primase (23); this site does not have a promoter. By analogy with mammalian mitochondrial replication, the model for initiation of yeast mtDNA replication proposes that second (non-promoter) strand synthesis is primed by a primase within the ori/rep region after first strand synthesis exposes the second strand origin (although, unlike mammalian mitochondria, the first and second strand origins are nearby in the ori/rep region). A stem and loop similar in secondary structure to the mammalian second strand origin can form between clusters A and B (6), but there is no evidence that second strand priming occurs at this site. The S1 nuclease technique was used to map a RNA 5′-end and a RNA–DNA transition for second strand synthesis (21); these sites are at the C cluster, as shown in Figure 1.

In this paper we re-examine the model for yeast mitochondrial replication by investigating ori-5 in the hypersuppressive petite HS40. Using the highly specific technique of *in vitro* capping with guanylyltransferase and [α-^{32}P]GTP, we labeled the initiating ribonucleotide triphosphate retained on the 5′-end of intact RNA primers of replication intermediates. This labeling distinguishes primed intermediates from the large excess of non-replicating DNA present in a restriction digest of mtDNA. *In vitro* capping has been used to identify the initiating ribonucleotide and RNA–DNA transition position on Okazaki fragments during discontinuous DNA synthesis in *Escherichia coli* (24) and *Drosophila melanogaster* (25). This technique was also previously used to identify the sites of transcription initiation for RNA from yeast mitochondria (8,17,26–30). Our results show that the promoter strand is indeed primed from the promoter, that the non-promoter strand origin is outside the conserved ori/rep region, and that for both strands the transition from RNA to DNA synthesis does not occur at a discrete site.

MATERIALS AND METHODS

Yeast strain and growth media

For these experiments we used HS40 (31), more specifically strain 14WWHS40 (MATα ρ$^-$ *ade2-1 ura3-52 trp1 leu2 CIT*::*LEU2*). Its mtDNA is virtually identical to the sequence between 18 837 and 19 605 in the GenBank sequence of yeast mtDNA (accession no. M62622; 32). The yeast were grown in medium consisting of 1.5% peptone, 1.5% yeast extract and 2% of the sugar raffinose, glucose or sucrose. The medium was supplemented with 25 mg/l adenine.

Mitochondrial DNA isolation

With minor modifications, we spheroplasted yeast by the method of Yaffe (33), except that we used two different enzymes, Yeast Lytic Enzyme (ICN) and Murienase (Amersham), both at 1 mg/g yeast. The spheroplasts were disrupted as described by Casey *et al*. (34) and then the mitochondria were isolated as described by Yaffe (33), except PMSF was omitted. For one experiment the mitochondria were quickly broken with glass beads using a Bead-Beater (Biospec Products). The mitochondrial pellets were lysed with 1% SDS and then the mtDNA was purified by two successive bis-benzimide/CsCl gradients (35). The mtDNA was ethanol precipitated, redissolved and repeatedly extracted with phenol and chloroform until the interface was clean, and finally ethanol precipitated.

Capping reactions

DNA for the capping reactions was generally cut first with a restriction enzyme. The digestion used a several-fold excess of both enzyme and time, was performed in the presence of 0.02 U/µl reaction Prime RNase Inhibitor (5′–3′ Inc.) and was stopped by phenol and chloroform extractions and ethanol precipitation. The DNA was dissolved in water and boiled for 3 min before adding the capping reaction ingredients. The capping reaction volume was usually 2 µl with 1–2 µg/µl DNA and the reaction was done with 50 mM Tris–HCl, pH 8.0, 6 mM KCl, 1.25 mM $MgCl_2$, 2.5 mM DTT, 100 µM *S*-adenosylmethionine, 100 µg/ml BSA, 6–10 µM [α-^{32}P]GTP (3000 Ci/mmol), 0.02 U/µl Prime RNase Inhibitor, 0.4 U/µl pyrophosphatase (Amersham) and 0.25 U/µl vaccinia guanylyltransferase (Gibco BRL). The reaction was incubated at 37°C for 90 min, then stopped by adding 48 µl water and purifying the product using the Qiaex II DNA extraction kit (Qiagen). The resin was eluted three times with 10 mM Tris–HCl, pH 8.0, 1 mM EDTA and the eluates pooled before ethanol precipitation.

Nuclease reactions

We used pancreatic deoxyribonuclease I (DNase I) digestion when we needed to demonstrate that DNA was part of capped nucleic acid molecules. Before stopping the capping, the 2 µl reaction was diluted 10-fold with 40 mM Tris–HCl, pH 7.8, 10 mM NaCl, 6 mM $MgCl_2$, 0.025 U/µl Prime RNase Inhibitor and 12 U RNase-free DNase I (Promega). After digestion at 37°C for 1 h 30 µl water was added and the nucleic acid purified using the Qiaex II kit. To completely remove the DNA portion of

primed replication intermediates gel-isolated capped fragments were digested as above with DNase I. Then the reaction was made 10 mM β-mercaptoethanol and 0.3 U T4 DNA polymerase (New England Biolabs) were added. After 1 h incubation at 37°C the reaction was phenol extracted, chloroform extracted and lyophilized to dryness. For RNA sequence analysis by the technique of partial digestion with base-specific ribonucleases we used the Amersham RNA sequencing kit. Limit T1 digestions were similar except that 250 U T1 ribonuclease were added.

Electrophoresis and fragment elution

Nucleic acid pellets were dissolved in loading buffer containing 80% formamide and 1× running buffer and boiled for 2 min before loading onto polyacrylamide gels. The gels were simply covered with plastic wrap for autoradiography, so that when fragments needed to be excised and eluted the gel was already hydrated. The autoradiographs were scanned with an imaging densitometer (BioRad Laboratories, GS-670). To elute the fragments the polyacrylamide pieces were crushed in a microfuge tube and suspended in 500 µl 0.5 M ammonium acetate, 10 mM Tris–HCl, pH 8.0, 1 mM EDTA and 0.1% SDS. After rotation at 37°C overnight the tubes were centrifuged and the eluate collected and counted in a scintillation counter. Elutions were repeated and pooled if necessary to free most of the label. The eluate was then filtered through a 0.45 µm micro spin filter, phenol extracted and chloroform extracted prior to ethanol precipitation in the presence of 1 µg carrier tRNA.

RESULTS

Mapping the priming sites of HS40

HS40 is a 757 bp hypersuppressive petite containing ori-5. To map its priming sites for DNA replication, we isolated the mtDNA and digested it with restriction enzymes which cut the petite genome once. The digested DNA consisted mostly of full-length duplex fragments but also included replicating fragments containing RNA primer. The nucleic acid of a replicating strand should extend from the primer to the restriction cut and this length will be discrete if the priming site is precise. Next, the DNA was denatured by boiling to make the duplexed RNA primer available for the *in vitro* capping reaction. The capping reaction will only label unprocessed RNA primers which retain a 5′-triphosphate. After capping the labeled RNA/DNA chimeric molecules were analyzed by gel electrophoresis and autoradiography.

A typical gel is shown in Figure 2A. When the HS40 DNA was cut with restriction enzymes, as in lanes 4–6, two capped bands are present. The most likely explanation for two bands would be two replication initiation sites, one for each DNA strand. In lane 3, where the DNA was cut with *Dra*I before capping, only one band is visible, but subsequent analysis (below) suggests that the band is a double band. In lane 1 HS40 mtDNA was not cut before the capping reaction. There are no bands in lane 1; label was incorporated but it is present as a background smear of random DNA lengths. While the restriction digestions showed that the capped products contain DNA, as would be expected for replication intermediates, the presence of DNA was further demonstrated by DNase I digestion. An equal aliquot of capped *Msp*I-cut DNA electrophoresed in lane 6 was digested with RNase-free DNase I and loaded in lane 7. The bands present in lane 6 disappeared, as did the background smear. We checked the specificity of the DNase I digestion by treating capped mitochondrial 15S rRNA from the petite P2 (36) under the same conditions; no degradation was visible after polyacrylamide gel electrophoresis (data not shown).

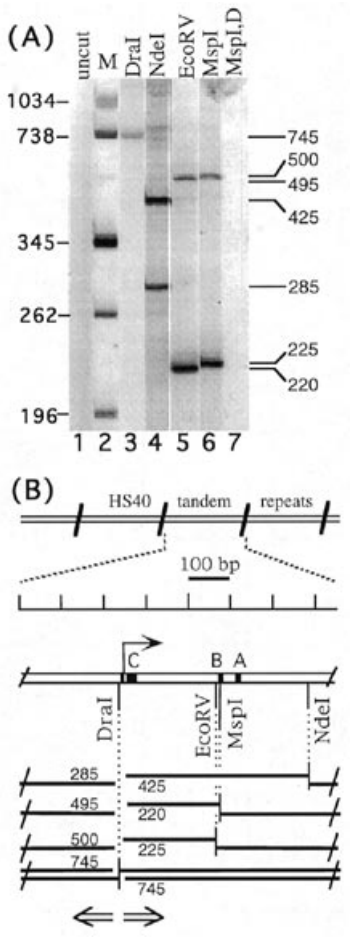

Figure 2. Restriction enzyme analysis of capped HS40 mtDNA electrophoresed on an 8% polyacrylamide–8 M urea gel. (**A**) Lane 2 has marker DNA (pBluescript II KS+), cut with *Sau*3AI and labeled by end-filling. Marker fragment sizes are given on the left side of the gel. All other lanes have capped HS40 mtDNA. In lane 1 the DNA was not cut with a restriction enzyme before capping. In lanes 3–7 the DNA was cut with various restriction enzymes before capping (lane 3 with *Dra*I, lane 4 with *Nde*I, lane 5 with *Eco*RV, lanes 6 and 7 with *Msp*I). After the capping reaction the DNA in lane 7 was digested with DNase I. In this figure lane 4 was replaced with a shorter exposure of the same lane. The sizes of the capped fragments were estimated from the postions of the marker bands using a graphing program (KaleidaGraph, Abelbeck Software); the capped fragment sizes and postions are marked on the right. (**B**) The physical map of the 757 bp repeat unit of HS40 mtDNA is shown with the positions of the three GC-rich clusters, C, B and A, and the four restriction sites. Diagonal lines indicate the boundary of the HS40 genome unit (the petite deletion junction); restriction fragments extend beyond the junction into the next tandem repeat. A bent arrow marks the ori-5 promoter. Below the map are four sets of horizontal bars, one set for each restriction enzyme. The bars show the map alignment and lengths of each enzyme's pair of capped fragments (from A). Each fragment is aligned from its cut end to its capped end. The two diverging arrows at the bottom of the figure indicate the priming sites identified by the capped ends of the fragments.

We estimated the lengths of the fragments in order to map the priming sites, since each discrete fragment presumably extends from a priming site (5′-end) to a restriction cut. This size information was used by aligning the fragments along the

restriction map, starting from the restriction sites. Any pair of fragments from a single lane could be oriented in two possible ways (i.e. the shorter one to the left of the restriction site and the longer to the right or vice versa). When the products of the *Nde*I, *Eco*RV and *Msp*I reactions were aligned using both possible orientations, one orientation of each set proved to have primed ends mapping in the vicinity of ends from the other enzymes (Fig. 2B), thereby identifying the two priming sites. One of these priming sites maps to the promoter. The capped band mapping to the promoter was always more intense than the other capped band, as can be seen in lanes 4–6 of Figure 2.

For the *Dra*I digestion a single capped band of 745 nt, nearly genome length, was visible on the polyacrylamide gel (Fig. 2, lane 3). Since the *Dra*I site is bracketed by the priming sites mapped with the other three enzymes, we believe that the single band is actually two bands of similar length extending around the HS40 genome, as shown by the last set of horizontal bars at the bottom of Figure 2.

Since replication intermediates may be short lived, we wondered if mtDNA from swiftly broken yeast would preserve more replication intermediates for capping. We isolated mtDNA from yeast quickly broken using glass beads and compared its capping efficiency with DNA isolated by enzymatic digestion of the cell wall. The intensity of the capped bands did not change (data not shown). We also considered the effect of growth stage of the yeast on the efficiency of capping. Yeast were harvested at various points during growth, from middle through late log and into early stationary phase. There proved to be no correlation between the intensity of capped bands and growth stage (data not shown).

We studied the effect of different culture media on capping reaction efficiency. We thought it likely that mtDNA isolated from derepressed mitochondria would cap more efficiently because derepressed mitochondria may have more mtDNA per cell and do have higher levels of mitochondrial gene expression (37). For this reason we had used the non-repressing sugar raffinose to grow the petite HS40 in most of our experiments (for example Fig. 2). We compared the capping efficiency of mtDNA isolated from yeast grown with the repressing sugar glucose or sucrose with the mtDNA from the yeast grown on raffinose. The same pair of capped bands were present, but their intensity was reduced (data not shown). In addition, the intensity of the capped band not mapping at the promoter, noted above as being the less intense of the two bands in a lane, became even less relative to the band mapping to the promoter.

Precise mapping of the 5′-end of the promoter strand primer

We used a conveniently located *Hph*I restriction site to show that the primed intermediate mapping to the promoter does in fact arise from the promoter initiation site. *Hph*I is a type IIS restriction enzyme, with the recognition and cleavage sites 5′-GGTGANNNNNNNN'N. The single *Hph*I site in HS40 is at the C cluster, so the enzyme respectively recognizes and cleaves the DNA 18–22 and 30 bp downstream of the promoter. However, *Hph*I should not cleave the fraction of HS40 replication intermediates whose RNA primers extend >17 nt, into the *Hph*I recognition sequence.

When HS40 was capped and cleaved with *Hph*I a band ~30 nt long appeared (Fig. 3, lower panel, lane 3), confirming that priming of DNA replication occurs at the promoter. This band is absent in lane 4, which was cut with another restriction enzyme.

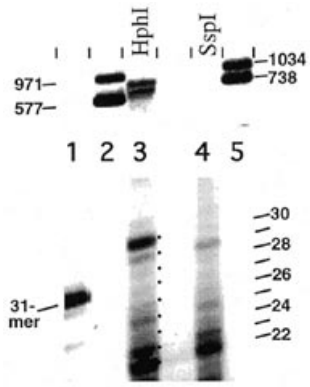

Figure 3. Accurate mapping of the promoter strand 5′-end. The two panels are sections from the top and bottom of the same 8% polyacrylamide gel, but the top panel is from an eight times longer exposure. DNA markers are in lanes 1, 2 and 5. Lane 1 has a kinase-labeled 31mer oligonucleotide. Lanes 2 and 5 have pUC18 and pBluescript II KS+ respectively, cut with *Sau*3AI and labeled by end-filling; sizes of fragments are given on the sides of the upper panel. Lane 3 has capped HS40 mtDNA cut with *Hph*I before capping and in lane 4 the DNA was cut with *Ssp*I. Sizes of the small capped RNA molecules are given on the right and the same bands are marked with dots by lane 3.

The 30 nt length is estimated from an end-labeled 31mer DNA oligonucleotide marker in lane 1; this 31mer reproducibly migrates with capped RNA of 27–28 nt length. (RNA, especially when capped, does not exactly co-migrate with single-stranded DNA.) Furthermore, the band migrates at a position 8 nt above a bright band previously shown to be 22 nt in length. The 22 nt band is part of a series of short RNA molecules which contaminate our mtDNA preparations (see Discussion for more about these short RNAs).

In the upper panel of Figure 3, lane 3, two bands of ~800 and 700 nt length are present. Non-promoter strand replication intermediates should extend from their start site to the *Hph*I site 694 nt away, providing the second band in lane 3. Promoter strand replication intermediates which did not cleave with *Hph*I at 30 nt, due to the presence of heteroduplexed primer RNA, should cut after the next tandem repeat. This will produce a fragment of 787 nt, which is seen as the top band in lane 3 of the upper panel of Figure 3. The 787 nt band is fainter than the 30 nt band, despite the longer exposure for the upper panel. This suggests that most of the replication intermediates have primers shorter than 18 nt.

RNA sequence and RNA–DNA transition for the promoter strand primer

Mapping capped restriction fragments from replication intermediates successfully localized priming sites. To identify the exact nucleotide positions for priming we sequenced the RNA ends of the primed intermediates. Appropriate capped bands from ~10 capping reactions, like those of Figure 2, lanes 4–6, were sliced out of autoradiographed gels and pooled. The eluted capped restriction fragments were subjected to enzymatic RNA sequencing reactions.

An example of sequence analysis for the capped intermediate which appears to originate from the promoter is shown in Figure 4. In lane 1 partial digestion with ribonuclease T1, which is specific for cleavage after guanosine (G) bases, produces a pattern of 5′…GGGGGXGGGGGXGGG…3′, where X is any base except G. In the HS40 genome only the sequence of cluster C,

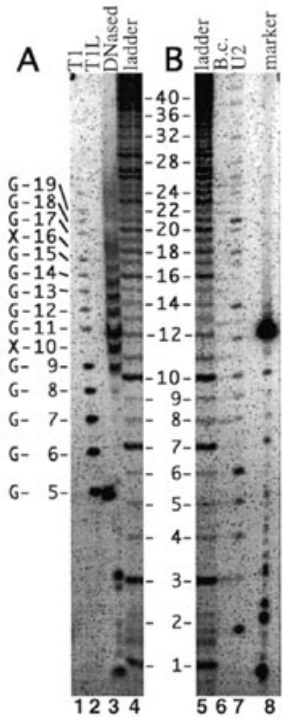

Figure 4. RNA sequencing and DNase treatment of replication intermediates. (**A**) Gel-isolated capped material arising from the promoter was subjected to partial T1 digestion in lane 1, a limit T1 digestion in lane 2 and a limit digestion with DNase I and T4 DNA polymerase in lane 3. The positions of the G residues in lanes 1 and 2 are marked on the left. Lanes 4 and 5 have size markers, a RNA ladder made by partial RNase PhyM (A and U-specific) digestion of capped petite P2 RNA (mostly consisting of 15 S rRNA, whose initial sequence is almost entirely A and U; see 36). The ladder nucleotide positions are marked between the two panels. (**B**) Gel-isolated capped material arising from the non-promoter strand was partially digested with *B.cereus* RNase in lane 6 and nuclease U2 in lane 7. Lane 8 is a size marker consisting of a limit T1 digestion of capped petite P2 RNA, whose predominant 15S rRNA has a G at position 12; this lane was overloaded to provide bands at other nucleotide positions. (A) and (B) are from the same autoradiogram of a 20% polyacrylamide–8 M urea gel.

downstream of the ori-5 promoter, matches this pattern of consecutive G residues. Limit digestion with T1 in lane 2 produces a single product which migrates at the same site as the first G in the sequence above. This G is therefore the first G in the RNA primer. The limit T1 ribonuclease digest has a doublet band caused by conversion of some of the 3′-2′ phosphate cyclic intermediates of the partial digestion into 3′-P. The RNA ladder of lane 4 shows that the first G is the fifth nucleotide of the primer. (Above nucleotide 5 the G-rich primer is retarded in migration relative to the RNA ladder, as can be seen in lanes 1 and 4. The retardation is reproducible. The ladder aligns well with other RNA, as seen in lanes 5–8.) The fact that the fifth nucleotide is the first G in the replication intermediate has been demonstrated by other means. The oligoribonucleotide from the limit digestion co-migrates with limit T1 digestion products from other RNA molecules with a G in their fifth positions. Furthermore, when the oligoribonucleotide was isolated and sequenced it proved to be a pentamer (data not shown). Replication intermediates therefore are primed from the promoter initiation site, as shown in Figure 1, starting at the same position as RNA transcripts (8).

We also investigated the location of the RNA–DNA transition by doing sequential digestions with RNase-free DNase I and T4 DNA polymerase (utilizing its 3′→5′ exonuclease activity). While digestion with either of these enzymes should remove the DNA from the chimeric replication intermediates and leave the RNA primer, we used two enzymes to further ensure complete digestion and to eliminate any concerns arising from the moderate base specificity of DNase I (38). We demonstrated the absence of RNases in this digestion by doing the same reaction with capped 15S rRNA from petite P2; the test RNA did not degrade (data not shown). A DNase I/T4 DNA polymerase digestion of the promoter-derived replication intermediates is in lane 3 of Figure 4. Base positions 1–5 in the DNase digestion can be determined from lane 4; beyond nucleotide 5 the base positions can be determined by comparison with the partial T1 digestion of the primer in lane 1. The most abundant DNase oligonucleotide product length is 11 nt; the large majority of the products range from 9 to 17 nt. A minority of oligonucleotides are longer, even beyond 30 nt. There are also some oligonucleotides 3 nt in length, but products with lengths from 4 to 8 nt are virtually absent. The progressive reduction in intensity of capped oligonucleotides longer than 10 nt in the partial T1 digest in lane 1 is probably due to the RNA–DNA transitions, which progressively reduce the substrate cleavable by T1.

RNA sequence and RNA–DNA transition for the non-promoter strand primer

In order to precisely determine the priming site for non-promoter strand replication intermediates, capped restriction fragments from these intermediates were isolated and partially digested with the pyrimidine-specific ribonuclease from *Bacillus cereus* (Fig. 4, lane 6) and the adenosine (A)-specific ribonuclease U2 (lane 7). When interpreting the sequence from lanes 6 and 7, if both lanes had a band at a position it was scored (U/A) and if bands were faint or absent in both lanes at a position it was scored N. As is often the case, the first nucleotide is absent from the digests. The RNA sequence of the non-promoter strand was read as follows: 5′-NA(U/A)AA-AN(U/A)AAUAUA(U/A)(U/A)AA(U/A)AA(U/A)(U/A)A(U/A)-ANANNNAN…-3′. This sequence aligns perfectly with the HS40 DNA sequence at the location where we first mapped the primer for the non-promoter strand. The first base of the primer for the non-promoter strand is 34 bp upstream of the first nucleotide of the primer for the promoter strand.

The RNA–DNA transition for the non-promoter strand replication intermediates was identified by DNase I/T4 DNA polymerase digestion (see Fig. 5, lane 1). The oligonucleotide products have lengths of 8, 11, 13, 16, 19, 23 and 25 nt and at every position in a region extending from 27 to 46 nt; a few are longer than 46 nt. Within the region 27–46 nt products with lengths of 30, 31 and 37–42 nt are somewhat more abundant.

DISCUSSION

For this study we used guanylyltransferase to cap replication intermediates present among restriction fragments from HS40 and demonstrated that each DNA strand has a single starting site for DNA synthesis. Sequence analysis showed that for both strands priming starts at a single discrete nucleotide, that the starting sites for the two strands are 34 bp apart and that synthesis is divergent from these starting sites (Fig. 6). The chimeric (RNA

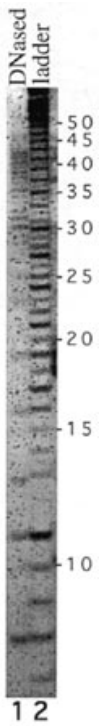

Figure 5. DNase treatment of replication intermediates from the non-promoter strand. Gel-isolated restriction fragments were digested with DNase I and T4 DNA polymerase in lane 1. Lane 2 contains a RNA ladder made from petite P2 RNA partially digested with RNase PhyM. The nucleotide lengths are marked on the right. The gel is 20% polyacrylamide–8 M urea.

and DNA) nature of the capped intermediates was demonstrated by the capping reaction itself, which only labels RNA with a 5′-tri- or diphosphate, and by DNase digestion. We identified the sites for RNA–DNA transition in the replication intermediates by digesting the capped replication intermediates with DNase I and T4 DNA polymerase. These enzymes leave a single deoxynucleotide attached to RNA when used to remove the DNA from a 5′-RNA–DNA-3′ chimeric molecule (24); for this reason the lengths of the primers summarized in Figure 6 are 1 nt shorter that the oligonucleotides measured in Figures 4 and 5.

In our experiments only a small fraction of the mtDNA was capped. This shows that replication intermediates, at least those with intact 5′-ends, are rare. We usually grew HS40 on the derepressing sugar raffinose because we found that mtDNA isolated from yeast grown on the repressing sugar glucose or sucrose capped less efficiently. Since we used similar amounts of DNA in the capping reactions, mtDNA from repressed yeast have a lower fraction of replication intermediates. This is probably due either to relatively faster turnover of replication intermediates or increased use of some other alternative replication pathway, like rolling circles (39). We also consistently found that the strand primed from the promoter capped better than the non-promoter strand (Fig. 2); this difference in efficiency was even more pronounced under repressing conditions. Since both strands of the DNA must replicate the same number of times, primers for the non-promoter strand must either turn over faster than primers for the promoter strand or the non-promoter strand must sometimes use an alternative priming mechanism, like discontinuous replication.

When we analyzed the strand primed at the promoter with DNase, we found that the vast majority of primers extended to the middle of the C cluster. The single most common primer length is 10 nt, most are 8–13 nt long and the frequency of longer primers falls gradually so that only a few are 28 nt or longer. There is also a 2 nt primer. The spot at the single nucleotide position in lane 3 of Figure 4 cannot be from a capped replication intermediate because the shortest possible DNase product must at minimum possess both a capped ribonucleotide and a deoxyribonucleotide. The fact that the majority of the promoter strand primers extend

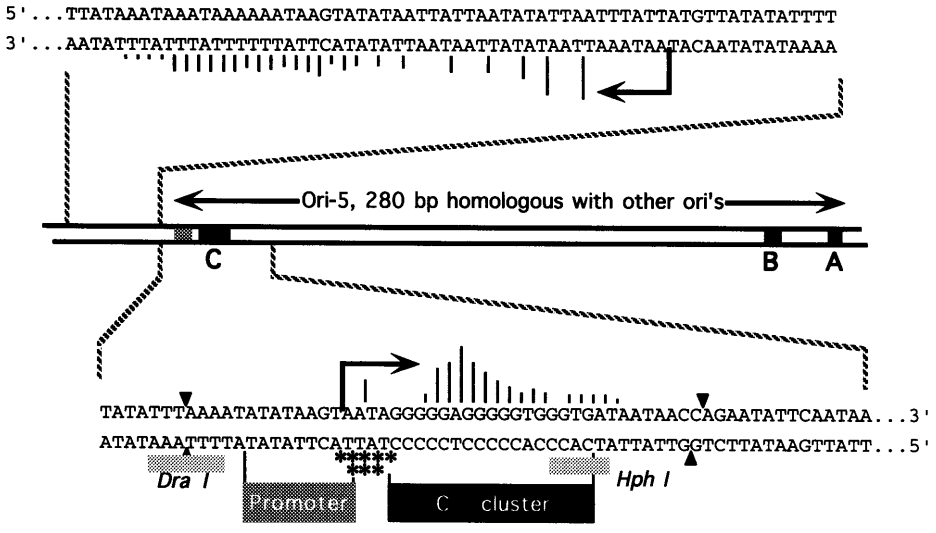

Figure 6. Replication start sites and primer lengths for ori-5. The top and bottom DNA sequences abut. The bent arrows mark the discrete 5′-ends of the primers for each strand, as determined in this paper; the bent arrow by the promoter is also the promoter transcription start site. The vertical lines above and below the sequence show the sites of RNA–DNA transitions; the relative length of the lines indicate the amount of primers of that length. Two restriction enzyme recognition sequences are marked with gray bars and their cleavage sites are marked with carets. Asterisks mark where the yeast MRP enzyme cleaves RNA *in vitro* (22).

longer than 5 nt is confirmed in Figure 4, where a limit T1 digestion of the replication intermediates produces a pentamer. This result would be impossible if the RNA–DNA transition was at or upstream of nucleotide 5. The fact that most of the primers are shorter than 18 nt is confirmed by *Hph*I cleavage (Fig. 3).

In Figure 3 a series of *in vitro* capped RNA molecules is present in lanes 3 and 4. When sequenced the RNAs proved to be from the HS40 ori-5 promoter initiation site, i.e. 5′-AAUAGGGGG… (data not shown). Most of these short RNAs in the HS40 mtDNA preparations can be removed by size fractionation. Some of the RNA persists unless the mtDNA is vigorously denatured, which suggest that it is heteroduplexed to the mtDNA. The short RNA molecules could simply be degraded transcripts. Alternatively, some of these RNAs could be made for priming promoter strand DNA replication. *In vitro* transcription from the yeast ori/rep promoter has been shown to produce a stable RNA–DNA heteroduplex which can prime *in vitro* DNA synthesis by yeast mtDNA polymerase (40). Clearly, all of the short *in vivo* RNAs cannot serve as primers, since the RNA is abundant and in most cases not duplexed. If any of the short RNA molecules are used as primers they must first be processed, because the RNA modal length is 22 nt while the modal length of the primers retained on replication intermediates is 10 nt.

The 3′-OH of an RNA primer must be generated by either cessation of synthesis or by cleavage from a longer precursor. In mammalian mitochondria MRP cleavage of RNA transcripts from the promoter may provide the primer 3′-OH (16,19,20), and a similar MRP is found in yeast (22). Yeast MRP cleaves at the right position to be responsible for the 2 nt primer we identified from the promoter strand DNA synthesis start site (Fig. 6), but the MRP cleavage site is not close to the rest of the RNA–DNA transitions. This suggests that yeast MRP has at best a minor role in processing primers. A nuclease from vertebrates, endonuclease G, cleaves RNA *in vitro* at poly(G) tracts and RNA from the CSB 2 region of animal mitochondria. Endonuclease G has been proposed to cleave primer RNA for mammalian mitochondrial replication (41). Since the majority of the promoter strand primers end in the G-rich cluster C, if a similar enzyme were present in yeast it could generate the 3′-ends of the most abundant promoter strand primers.

The priming site for the non-promoter strand is 34 bp upstream of the promoter primer site, outside the conserved ori/rep region entirely. We found by DNase analysis that the zone of the RNA–DNA transition for this strand is even larger than for the promoter strand. The shortest primer is 7 nt and a few extend beyond 45 nt. The various primers with lengths within the range 7–26 nt share a common feature, that their 3′-ends occur immediately before the sequence AT on the template strand (Fig. 6). Longer primers end at every position from 26 to 45 nt; this cluster of primer lengths is centered on a stretch of six A residues on the template strand. Because of the sequence differences, it is unlikely that the 3′-OH ends of the non-promoter strand primers are produced by the same mechanism as the promoter strand.

The selectivity of the *in vitro* capping technique gives us confidence in our conclusions, but Baldacci *et al.* (21) present different results for ori-5 (Fig. 1). They used another petite, but that is probably not the cause of the differing results. While HS40 does have a genome ~200 bp longer than the petite they used, the shared DNA is virtually identical, with the same ori/rep and flanking sequences. The difference may arise because they used pancreatic RNase A to digest some of the mtDNA before S1 analysis in order to compare the 5′-ends of replication intermediates with and without attached primers. RNase A is pyrimidine-specific, so in the case of the purine (mostly G)-rich beginning of the promoter strand, RNase A could cleave after nucleotide 3, a U, then fail to remove the rest of the primer RNA. Within the resolution of the S1 technique this could explain why they found a primer with a length of 4 nt. (Alternatively, their 4 nt primer could be the same as our minor 2 nt primer, again within the limits of S1 resolution.) There is also no mention of denaturing the mtDNA before their RNase A digestion (21), so digestion of RNA primers from the duplex would be poor at best.

Our confirmation of primer synthesis from the initiation site of the ori/rep promoter supports the premise in the model outlined in the Introduction that mtRNA polymerase (consisting of the core polymerase coded by the *RPO41* gene and the specificity factor mtf1) synthesizes the primer. Prior genetic evidence for this idea comes from the fact that mitochondrial ori/rep regions and their associated promoters are required for hypersuppressitivity (2,3,7,17), although not needed for petite mtDNA stability (42,43), and the fact that disruption of *RPO41* in wild-type yeast caused the mtDNA to became unstable, producing rho° petites (petites lacking mtDNA) (44). However, Fangman *et al.* (43) showed that both rho⁻ and rho° petites were produced. Furthermore, the mtDNA of established rho⁻ yeast, with or without an ori/rep region, proved to be stable with such disruptions and, significantly, *RPO41* disruption did not affect petite hypersuppressitivity (4). Disruption of *RPO41* may simply cause the mtDNA instability by preventing mitochondrial protein synthesis, which is known to produce petites (37). At this point the actual role of core mitochondrial RNA polymerase in priming is not clear. When present, the core mtRNA polymerase may be used to synthesize the promoter strand primer, while an alternative pathway may be used when the *RPO41* gene product is absent. On the other hand, core polymerase might actually never be used to prime DNA synthesis.

Since core polymerase is not necessary for hypersuppressitivity but the promoter is required, the mitochondrial transcription specificity factor mtf1 (also called sc-mtTFB) may play a role in priming. Together with the core polymerase mft1 recognizes the promoter sequence during transcription (45,46). If core polymerase does indeed play a role in priming, mtf1 surely also plays its role in recognizing the promoter; if core polymerase has no priming role or is absent, mtf1 might work with an alternative enzyme like primase to recognize the promoter. Current genetic evidence for the replication role of mtf1 is ambiguous. Yeast with a disrupted *MTF1* gene become rho° (47), but unfortunately the effect of this disruption on established rho⁻ petites was not reported. Temperature-sensitive *MTF1* mutants did not lose mtDNA for several generations and petites did not appear until after 24 h (48), suggesting that mtf1 is not directly required for all mtDNA replication.

Beyond the fact that promoter strand replication is primed from the promoter, our results do not support a closely analogous mechanism for yeast and mammalian mitochondria. For ori-5 replication the starting sites for the two strands diverge and there is no stem and loop at the non-promoter strand starting site. There is also no evidence for, or against, synthesis at one site starting before the other or that the DNA needs to be single-stranded at one of the starting sites.

We have shown that the non-promoter strand starting site is located outside the conserved ori/rep sequence. Our identification of one non-promoter priming site does not tell us the sequence signal for priming. Computer searches using various 10 bp DNA sequences next to or centered on the non-promoter strand start site found many matches throughout the genome, which is unsurprising considering the high AT content of the genome and the search sequences. Significantly, matches were not found at similar positions relative to the promoter in other ori/rep regions. In the middle of the 34 bp between the two start sites is the sequence TTTTATATTTA, which was pointed out after comparative sequence analysis of the ori/rep regions (6) to be identical to a yeast nuclear ARS site. Related sequences, named *r*a sequences, are found upstream of cluster C in most of the ori/rep regions (6).

A probable mtDNA packaging protein variously called ABF2, p19/HM or sc-mtTFA (11) has been shown to specifically footprint in the same region between the start sites, upstream of the ori-5 promoter sequence (49). Its 25–30 bp footprint (50) therefore fills the space between the non-promoter strand start site and the promoter and could serve to recruit primase for non-promoter strand replication. Disruption of the gene for ABF2 caused yeast grown non-selectively on glucose to become rho°, although with glycerol they grew slowly and the mtDNA was maintained.

An alternative to a sequence signal near the non-promoter strand priming site itself could be provided by proximity to the promoter. In this case the binding of mtf1 or some other factor(s) to the promoter could orient the molecules involved in priming to both starting sites. Footprinting experiments on the transcription initiation complex at yeast mitochondrial promoters show that the mitochondrial core RNA polymerase and the specificity factor mtf1 protect ~15 bp upstream from the promoter start (45), which is only 20 bp away from the non-promoter strand starting site. This footprint by mtf1 and core polymerase or a footprint by mtf1 and a primase would be well placed to guide another primase molecule and perhaps other factors to the non-promoter strand start site. If this hypothesis is true, in other hypersuppressive petites from other ori/rep regions the non-promoter strand start sites will also be located ~34 bp away from the promoter start site.

In this paper we show that *in vitro* capping combined with restriction digestion is useful for mapping origins of replication. The selectivity and power of the technique arises from the fact that only intact 5′-ends of primers can be labeled. The terminal label from the *in vitro* capping permits sequencing and determination of the length of the primers.

ACKNOWLEDGEMENTS

We thank Ronald Butow (Department of Biochemistery, University of Texas South Medical Center, Dallas, TX) for generously providing the petite HS40. This work was supported by the US Public Health Service, NIH grant GM46009.

REFERENCES

1 Dujon,B. (1981) In Strathern,J.N., Jones,E.W. and Broach,J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2 Blanc,H. and Dujon,B. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3942–3946.
3 de Zamaroczy,M., Marotta,R., Faugeron-Fonty,G., Goursot,R., Mangin,M., Baldacci,G. and Bernardi,G. (1981) *Nature*, **292**, 75–78.
4 Lorimer,H.E., Brewer,B.J. and Fangman,W.L. (1995) *Mol. Cell. Biol.*, **15**, 4803–4809.
5 Lockshon,D., Zweifel,S.G., Freeman-Cook,L.L., Lorimer,H.E., Brewer,B.J. and Fangman,W.L. (1995) *Cell*, **81**, 947–955.
6 de Zamaroczy,M., Faugeron-Fonty,G., Baldacci,G., Goursot,R. and Bernardi,G. (1984) *Gene*, **32**, 439–457.
7 Baldacci,G. and Bernardi,G. (1982) *EMBO J.*, **1**, 987–994.
8 Christianson,T. and Rabinowitz,M. (1983) *J. Biol. Chem.*, **258**, 14025–14033.
9 Walberg,M.W. and Clayton,D.A. (1981) *Nucleic Acids Res.*, **9**, 5411–5421.
10 Schinkel,A.H. and Tabak,H.F. (1989) *Trends Genet.*, **5**, 149–154.
11 Schmitt,M.E. and Clayton,D.A. (1993) *Curr. Opin. Genet. Dev.*, **3**, 769–774.
12 Clayton,D.A. (1982) *Cell*, **28**, 693–705.
13 Chang,D.D. and Clayton,D.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 351–355.
14 Hixson,J.E., Wong,T.W. and Clayton,D.A. (1986) *J. Biol. Chem.*, **261**, 2384–2390.
15 Clayton,D.A. (1987) *Phil. Trans. R. Soc. Lond. B*, **317**, 473–482.
16 Chang,D.D. and Clayton,D.A. (1987) *EMBO J.*, **6**, 409–417.
17 Osinga,K.A., De Hann,M., Christianson,T. and Tabak,H.F. (1982) *Nucleic Acids Res.*, **10**, 7993–8006.
18 Chang,D.D., Hauswirth,W.W. and Clayton,D.A. (1985) *EMBO J.*, **4**, 1559–1567.
19 Lee,D.Y. and Clayton,D.A. (1997) *Genes Dev.*, **11**, 582–592.
20 Bennett,J.L. and Clayton,D.A. (1990) *Mol. Cell. Biol.*, **10**, 2191–2201.
21 Baldacci,G., Cherif-Zahar,B. and Bernardi,G. (1984) *EMBO J.*, **3**, 2115–2120.
22 Stohl,L.L. and Clayton,D.A. (1992) *Mol. Cell. Biol.*, **12**, 2561–2569.
23 Wong,T.W. and Clayton,D.A. (1985) *Cell*, **42**, 951–958.
24 Kitani,T., Yoda,K. and Okazaki,T. (1984) *Mol. Cell. Biol.*, **4,** 1591–1596.
25 Kitani,T., Yoda,K., Ogawa,T. and Okazaki,T. (1985) *J. Mol. Biol.*, **184**, 45–52.
26 Levens,D., Ticho,B., Ackerman,E. and Rabinowitz,M. (1981) *J. Biol. Chem.*, **256**, 5226–5232.
27 Christianson,T., Edwards,J.C., Mueller,D.M. and Rabinowitz,M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5564–5568.
28 Edwards,J.C., Osinga,K.A., Christianson,T., Hensgens,L.A.M., Janssens,P.M., Rabinowitz,M. and Tabak,H.F. (1983) *Nucleic Acids Res.*, **23**, 8269–8282.
29 Martin,N.C., Miller,D.L., Underbrink,K. and Ming,X. (1985) *J. Biol. Chem.*, **260**, 1479–1483.
30 Backer,J.S. and Getz,G.S. (1987) *Nucleic Acids Res.*, **15**, 9309–9324.
31 Parikh,V.S., Morgan,M.M., Scott,R., Clements,L.C. and Butow,R.A. (1987) *Science*, **235**, 576–580.
32 de Zamaroczy,M. and Bernardi,G. (1986) *Gene*, **47**, 155–177.
33 Yaffe,M.P. (1991) *Methods Enzymol.*, **194**, 627–643.
34 Casey,J., Cohen,M., Rabinowitz,M., Fukuhara,H. and Getz,G.S. (1972) *J. Mol. Biol.*, **63**, 431–440.
35 Hudspeth,M.E.S., Shumard,D.S., Tatti,K.M. and Grossman,L.I. (1980) *Biochim. Biophys. Acta*, **610**, 221–228.
36 Christianson,T., Edwards,J., Levens,D., Locker,J. and Rabinowitz,M. (1982) *J. Biol. Chem.*, **257**, 6494–6500.
37 Costanzo,M.C. and Fox,T.D. (1990) *Annu. Rev. Genet.*, **24**, 91–113.
38 Bernardi,A., Gaillard,C. and Bernardi,G. (1975) *Eur. J. Biochem.*, **52**, 451–457.
39 Maleszka,R., Skelly,P.J. and Clark-Walker,G.D. (1991) *EMBO J.*, **10**, 3923–3929.
40 Xu,B. and Clayton,D.A. (1995) *Mol. Cell. Biol.*, **15**, 580–589.
41 Cote,J. and Ruiz-Carrillo,A. (1993) *Science*, **261**, 765–769.
42 Fangman,W.L., Henly,J.W., Churchill,G. and Brewer,B.J. (1989) *Mol. Cell. Biol.*, **9**, 1917–1921.
43 Fangman,W.L., Henly,J.W. and Brewer,B.J. (1990) *Mol. Cell. Biol.*, **10**, 10–15.
44 Greenleaf,A.L., Kelly,J.L. and Lehman,I.R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3391–3394.
45 Schinkel,A.H., Groot Koerkamp,M.J.A. and Tabak,H.F. (1988) *EMBO J.*, **7**, 3255–3262.
46 Shadel,G.S. and Clayton,D.A. (1993) *J. Biol. Chem.*, **268**, 16083–16086.
47 Lisowsky,T. and Michaelis,G. (1988) *Mol. Gen. Genet.*, **214**, 218–223.
48 Cliften,P.F., Park,J., Davis,B.P., Jang,S. and Jaehning,J.A. (1997) *Genes Dev.*, **11**, 2897–2909.
49 Diffley,J.F.X. and Stillman,B. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 7864–7868.
50 Diffley,J.F.X. and Stillman,B. (1992) *J. Biol. Chem.*, **267**, 3368–3374.