MOLECULAR AND CELLULAR BIOLOGY, Oct. 1985, p. 2770–2780
0270-7306/85/102770-11$02.00/0

Vol. 5, No. 10

Signals for Transcription Initiation and Termination in the *Saccharomyces cerevisiae* Plasmid 2μm Circle

ANN SUTTON† AND JAMES R. BROACH*

Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544

Received 26 April 1985/Accepted 5 July 1985

By S1 nuclease protection experiments and primer extension analysis, we determined precisely the cap and polyadenylation sites of transcripts from the four genes of the yeast 2μm circle plasmid, as well as those of other plasmid transcripts of unknown function. In addition, we used deletion analysis to identify sequences necessary for polyadenylation in plasmid transcripts. Our results indicate that plasmid genes constitute independent transcription units and that plasmid mRNAs are not derived by extensive processing of precursor transcripts. In addition, we found that the D coding region of 2μm circle is precisely encompassed by a polyadenylated transcript, suggesting that this coding region constitutes a functional plasmid gene. Our identification of the position of plasmid polyadenylation sites and of sequences necessary for polyadenylation provides support for a tripartite signal for polyadenylation as proposed by Zaret and Sherman (K. S. Zaret and F. Sherman, Cell 28:563–573, 1982). Finally, these data highlight salient features of the transcriptional regulatory circuitry that underlies the control of plasmid maintenance in the cell.

The yeast 2μm circle plasmid is a 6,318-base-pair (bp) double-stranded DNA species present in most *Saccharomyces cerevisiae* strains at 60 to 100 copies per haploid genome (4). The plasmid confers no apparent phenotype to cells in which it is resident. Nonetheless, the plasmid genome encodes products that are responsible for a number of activities that ensure maintenance of the plasmid. These activities include partitioning of plasmid molecules between mother and daughter cells after mitosis, plasmid amplification when copy number is low, and regulation of plasmid copy levels (20–22, 32, 36).

Recent genetic analysis has provided a correlation between specific coding domains within the plasmid genome and specific plasmid functions (Fig. 1). Three loci are required for stable propagation of the plasmid. These include two *trans*-active loci, *REP1* and *REP2*, which correspond to open coding regions B and C, and a *cis*-active locus, *REP3*, which consists of a series of direct tandem repeats of a 62-bp sequence. These three loci are responsible for ensuring plasmid partitioning and may also be involved in copy control (20–22). The product of the *FLP* gene, which corresponds to coding region A, catalyzes intra- and intermolecular recombination at a specific site within the inverted repeats of the plasmid (6, 31). Recent evidence suggests that this recombination is required to promote amplification of the plasmid (A. W. Murray and J. W. Szostak, personal communication). A fourth coding region, D, is evident from the nucleotide sequence of the plasmid, although no plasmid dysfunction or phenotype has been associated with disruption of this locus.

Consistent with this genetic analysis, our previous studies on 2μm circle transcription indicate that all three major open coding regions are transcribed in vivo (5). A transcript corresponding to the D region, however, was not unambiguously identified. A number of other polyadenylated [poly(A)$^+$] transcripts complementary to various regions of 2μm circle were also noted, several of which are indicated on the transcription map of the plasmid shown in Fig. 1. In fact, as is evident from Fig. 1, except for the region between the origin of replication and the *REP3* locus, the entire 2μm circle genome is transcribed into at least one poly(A)$^+$ species and often more. On the basis of the pattern of these transcripts, we proposed a model of 2μm circle expression that incorporated the processing of larger transcripts into functional mRNA species (5).

In this paper we report the precise locations of the 5′ and 3′ termini of the major transcripts of 2μm circle. In addition, we identified the transcriptional consequences of deletions near the 3′ ends of various plasmid genes. These results allowed us to address the question of whether some transcripts are derived from others. In addition, these data provide useful sequence information to help clarify the salient features of signals for transcription initiation and for polyadenylation and termination in yeasts. Finally, these data lay the groundwork for analyzing the intricate regulation of 2μm circle transcription that undoubtedly underlies the copy control system of the plasmid.

MATERIALS AND METHODS

Strains and plasmids. *Escherichia coli* C600 (*thr-1 thi-1 leuB6 supE44 lacY1 tonA21*) was used for propagation and amplification of hybrid plasmids. *S. cerevisiae* DCO4 (*MATa adel leu2-04*), either lacking endogenous plasmids or harboring an authentic 2μm circle plasmid or an appropriate hybrid 2μm circle plasmid, was used as our source of cellular RNA.

The hybrid plasmid pBR2μA, which consists of the entire 2μm circle genome (A form) cloned at the *Eco*RI site of the small unique region into the *Eco*RI site of pBR322, was the source of the DNA fragments used as probes for the transcript mapping. Plasmid CV20 and the Xho plasmid series derived therefrom have been described previously (20).

Nucleic acid preparation. Plasmid DNA was isolated from *E. coli* by a modification of the procedure of Birnboim and Doly (3) as described by Maniatis et al. (29). Poly(A)$^+$ RNA was obtained from yeast strains grown in either YEPD (5) (for transcript mapping) or SC-leucine (6) (for Northern analysis) medium as previously described (5).

* Corresponding author.
† Present address: Department of Biology, Brookhaven National Laboratory, Upton, NY 11973.

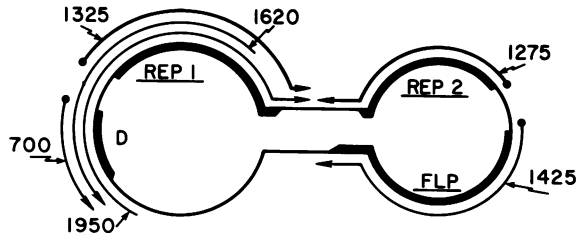

FIG. 1. Transcription map of the yeast 2μm circle plasmid. The map positions of the major 2μm circle transcripts are indicated on a diagram of the B form of the 2μm circle genome. Each transcript is designated by its length in bases, as determined by previous Northern analysis of 2μm circle transcription (5). The precise location of the 3′ end of all the transcripts shown (indicated by the arrows) and the 5′ ends of all but the 1,950- and 1620-base transcripts are reported in the text. The locations of the four open coding regions are indicated by the heavy lines on the diagram of the genome, with tapers lying at the 3′ end of the gene.

For preparation of end-labeled single-strand probes, plasmid pBR2μA was digested with the appropriate restriction enzymes (New England BioLabs, Inc., or Bethesda Research Laboratories, Inc.) as recommended by the supplier. The 5′ or 3′ ends of the DNA were labeled with ^{32}P as described below, and the appropriate fragment was purified by polyacrylamide gel electrophoresis. DNA strands were then separated by gel electrophoresis as described by Maxam and Gilbert (30). In certain cases, the probes were isolated from digests of larger restriction fragments rather than from the entire pBR2μA plasmid. In these cases, the larger fragments were purified by electrophoresis on agarose gels, electroeluted (30), and chromatographed on NACS-52 columns (Bethesda Research Laboratories) as recommended by the supplier.

DNA was labeled at its 5′ ends by using [^{32}P]ATP and T4 polynucleotide kinase (30) after dephosphorylation by treatment with calf intestinal alkaline phosphatase (29). The 3′ ends were labeled in restriction enzyme buffer by using the Klenow fragment of DNA polymerase and [α-^{32}P]deoxyribonucleoside triphosphates as described by Maniatis et al. (29).

Uniformly labeled single-strand probe for Northern analysis was prepared as follows. In plasmid pJR-0.9, the 1,314-bp *Hin*dIII fragment from the 2μm circle, which lies almost entirely within the *REP1* gene, has been cloned into the *Hin*dIII site of pBR322 and is thus flanked on one side by a unique *Eco*RI site and on the other side by a unique *Bam*HI site. Accordingly, plasmid pJR-0.9 DNA (1 μg) was digested with either *Bam*HI or *Eco*RI and then treated with a titrated amount of exonuclease III (12) to yield resection of approximately 1,500 bases from each end. After the resected DNA was purified with phenol and ethanol precipitation, the single-stranded regions were filled in by using the Klenow fragment of *E. coli* DNA polymerase I and [α-^{32}P]deoxyribonucleoside triphosphates. The sample was then digested with *Hin*dIII, and the 1,314-bp fragment was purified by electrophoresis on an agarose gel. The labeled fragment was then used as the hybridization probe in Northern analysis as described previously for nick-translated probes (5).

S1 nuclease mapping. S1 nuclease mapping was done by a modification of the method of Berk and Sharp (2). The single-strand DNA probe (10^4 to 10^5 cpm) was precipitated in ethanol with 25 to 50 μg of poly(A)$^+$ RNA isolated from [cir$^+$] and [cir^0] yeast strains. The precipitates were washed with 70% ethanol, dried, and suspended in 25 μl of hybrid-

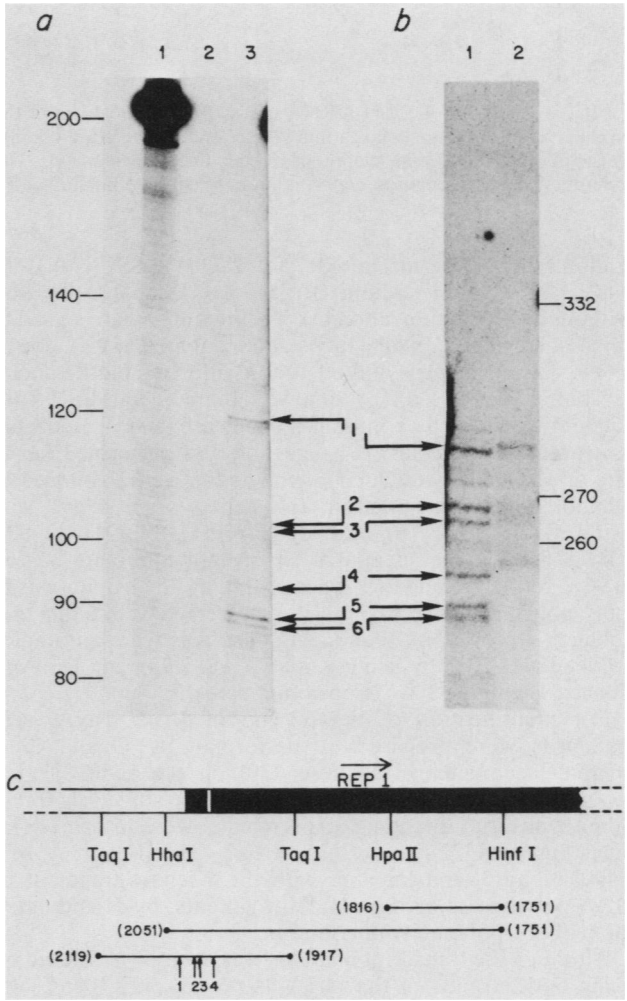

FIG. 2. Localization of the 5′ end of the *REP1* transcript by S1 nuclease digestion and by primer extension. For each sample, 25 μg of poly(A)$^+$ RNA from strain DCO4 [cir$^+$] or DCO4 [cir^0] was hybridized to the separated strands of one of the DNA fragments indicated in panel c as described in Materials and Methods. The two separated strands of each DNA fragment were used individually as probes. Results are shown only for that strand of the probe for which specific protection was observed. (a) S1 nuclease analysis with the *Taq*I fragment (position 1917 to 2119) as probe. Lanes: 1, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 2, probe hybridized to [cir^0] RNA, treated with 100 U of S1 nuclease; 3, probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added. Arrows indicate the major fragments protected from S1 nuclease digestion by *REP1* RNA. The numbers reflect fragment sizes in base pairs determined by fractionation of A+G and C+T sequencing ladders of the probe in the same gel (data not shown). (b) Comparison of 5′-end determination by S1 nuclease and primer extension analysis. Lane 1: The *Hpa*II-*Hin*fI fragment indicated in panel c was hybridized to [cir$^+$] RNA and used as the primer for cDNA synthesis by avian myeloblastosis virus reverse transcriptase as described in Materials and Methods. Lane 2: S1 nuclease analysis with the *Hha*I-*Hin*fI fragment indicated in panel c as the probe. The probe was hybridized to [cir$^+$] RNA and treated with 100 U of S1 nuclease. The numbers reflect the size of fragments (in base pairs) of pBR322 DNA digested with *Hha*I, end labeled, and fractionated on the same gel (data not shown). (c) Location of the probes and primer from the 5′ end of the *REP1* gene. The endpoints of the fragments are identified by the numbering system of Hartley and Donelson (13) for the A form of the 2μm circle plasmid. The arrows indicate the positions of the four groups of termini identified in panels a and b.

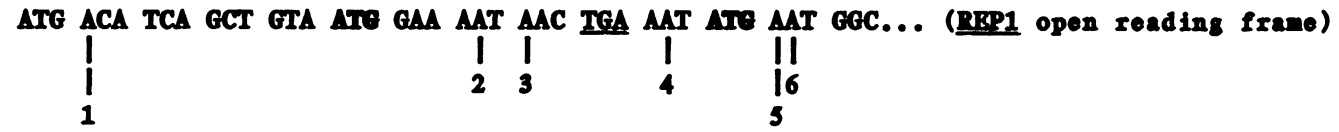

FIG. 3. Position of *REP1* mRNA cap sites. The locations of the six cap sites identified in the experiments shown in Fig. 2 are indicated on the nucleotide sequence spanning the 5′ end of the *REP1* coding region. The sequence shown extends from position 2041 to position 2000, by the plasmid numbering system of Hartley and Donelson (13). The potential in-frame translation initiation codons are in boldfaced type, and the intervening termination codon is underlined. The positions of the major transcript cap sites are marked by double vertical bars.

ization buffer (50% formamide, 5× SSC [1× SSC is 0.15 M NaCl plus 0.15 M sodium citrate], 1× Denhardt solution without bovine serum albumin). The mixtures were sealed in siliconized glass capillary tubes and incubated at 42°C for 16 h. The reactions were diluted 10-fold into ice-cold S1 nuclease buffer (4 mM $ZnSO_4$, 30 mM sodium acetate [pH 4.6], 250 mM NaCl) with or without S1 nuclease (100 U; Bethesda Research Laboratories). S1 digestion was performed for 45 min at 37°C and terminated by the addition of 2 volumes of ethanol. After precipitation, the pellets were dried, suspended in 80% formamide–10 mM NaOH–1 mM EDTA–0.1% xylene cyanol–0.1% bromphenol blue, heated to 90°C for 3 min, and then fractionated by electrophoresis in 8.0% polyacrylamide-urea gels as described by Maxam and Gilbert (30) for DNA sequencing. The gels were autoradiographed at −70°C by using Kodak XAR-5 film and Du Pont Cronex Lightning-Plus intensifying screens.

For identification of the 5′ end of *REP1* RNA, size standards were prepared by performing A+G and C+T chemical sequencing reactions (30) on the same labeled strand as was used for the probe. DNA size standards for all other transcript mapping experiments were prepared by digesting pBR322 with *Hha*I, *Hin*fI, *Hae*III, or *Hpa*II, followed by 3′-end labeling with the Klenow fragment of DNA polymerase or, for *Hae*III fragments, by 5′-end labeling with T4 polynucleotide kinase.

Although the transcript mapping was always performed by using both strands of the DNA as probe, only the results with the strand that was protected from digestion by [cir^+] RNA are shown.

Transcript mapping by primer extension. Primer extension to map the 5′ end of *REP1* mRNA was performed essentially as described by Ghosh et al. (11). Poly(A)$^+$ yeast RNA was hybridized to a 65-base *Hpa*II-*Hin*fI single-strand end-labeled DNA fragment as described above for S1 nuclease analysis. After 16 h at 42°C, the mixture was diluted into 200 μl of buffer (50 mM Tris [pH 8.1], 6 mM magnesium acetate, 60 mM NaCl, 10 mM dithiothreitol, 1 mM of each deoxyribonucleoside triphosphate [P-L Biochemicals, Inc.]) containing 13 U of avian myeloblastosis virus reverse transcriptase (Bethesda Research Laboratories). The reaction was incubated at 41°C for 3 h. NaOH was added to 0.2 M, and incubation was continued for 1 h. After neutralization with HCl, the mixture was extracted twice with phenol, precipitated twice with ethanol, and prepared for electrophoresis as described above for S1 mapping.

Northern analysis of 2μm circle transcripts. RNA was isolated from yeast strains grown in SC-leucine, poly(A) selected, fractionated on 2% agarose gels containing methyl mercury, transferred to diazotized paper, hybridized, and autoradiographed as previously described (5).

RESULTS

Identification of *REP1* cap site. In our initial localization of the *REP1* cap site, we used as probe a *Taq*I fragment (Fig. 2c) that spans the 5′ end of the coding region. We labeled the fragment at its 5′ ends with T4 polynucleotide kinase and [γ-^{32}P]ATP and then hybridized separated strands of the probe to poly(A)$^+$ RNA isolated from [cir^0] or [cir^+] yeast strains, that is, from strains that either lack or contain endogenous 2μm circles. We digested residual single-stranded nucleic acid with S1 nuclease and then determined the size of the labeled fragment protected by *REP1*-specific RNA by electrophoretic fractionation of the S1-resistant nucleic acid on denaturing polyacrylamide gels. An autoradiogram of one such fractionation is shown in Fig. 2a. With RNA from a [cir^+] strain, we observed considerable heterogeneity in the RNA-protected fragments (lane 3). DNA bands corresponding to two major and several minor cap sites span a distance of 30 bp at the 5′ end of the *REP1* coding region. We obtained no protection of the DNA probe from S1 digestion of labeled probe hybridized with RNA from a [cir^0] strain (lane 2). Thus, we concluded that the protected fragments seen in lane 3 derive from 2μm circle specific RNA.

To confirm the identification of the *REP1* cap sites obtained by S1 nuclease protection, we localized the 5′ ends of *REP1* transcripts by primer extension procedures. We used as primer a 65-bp *Hin*fI-*Hpa*II fragment that lies within the *REP1* coding region (Fig. 2c). We labeled the fragment at its 5′ ends, separated the strands of the fragment, and then hybridized the strands individually to freshly prepared poly(A)$^+$ RNA. We then used the hybridization mixture for reverse transcriptase-directed synthesis of DNA complementary to the hybridized RNA. As a direct comparison, we repeated S1 analysis of *REP1* transcripts, but in this case using a single-strand probe with the same 5′ terminus as the *Hin*fI-*Hpa*II primer fragment (Fig. 2c). For both sets of reactions, we determined the sizes of the synthesized or protected fragments by electrophoretic fractionation on denaturing polyacrylamide gels. As is evident from Fig. 2b, lanes 1 and 2, the sizes of the longest molecules synthesized by primer extension were identical to those of DNA fragments protected from S1 nuclease digestion. These results showed that *REP1* transcripts are not spliced within this region and confirmed the heterogeneity of *REP1* RNA cap sites.

We have indicated in Fig. 3 the locations of the *REP1* cap sites within the DNA sequence at the 5′ end of *REP1*. The 5′ end of the longest major transcript corresponds to position 2035 (2μm circle A form; 13), 9 bp upstream from an ATG in frame with the *REP1* coding region. However, this ATG is followed, after three codons, by a TGA translation terminator. The first ATG of the extended reading frame lies 3 bp downstream from the terminator. Several transcripts have cap sites lying between these two ATG codons. The cap sites of the remaining transcripts lie downstream from both codons and about 39 bp upstream from the next ATG, which is also in frame with the coding region. The significance of this unusual pattern of transcript cap sites is not known.

Identification of *REP2* and *FLP* cap sites. The initial ATG codons of the open coding regions for *REP2* and *FLP* lie 370 bp apart in the small unique region of 2μm circle. These genes are divergently transcribed and terminate within opposite inverted repeats (Fig. 1). To determine the locations of the cap sites of *REP2* and *FLP* RNAs, we used an S1 nuclease protection procedure analogous to that described above for the *REP1* gene. We used a 296-bp *Asu*I fragment extending from position 5029 to 5325 as probe for *REP2* RNA (Fig. 4c). Results from this S1 nuclease analysis are shown in Fig. 4a. As we observed for *REP1* transcripts, there appear to be multiple cap sites for *REP2* transcripts. Minor cap sites lie at positions 5223 and 5212, 25 and 14 bp upstream from the *REP2* open reading frame. Major protected fragments correspond to transcripts with 5′ ends at positions 5201 to 5198 and 5186 to 5184, that is, just upstream from the expected initiator ATG codon of *REP2* (position 5198 to 5196), or within the coding region itself.

The open coding region for *FLP* begins at position 5570 and ends at position 521. We used a *Taq*I fragment extending from position 5484 to 5779 as probe for the *FLP* cap site. S1 protection analysis with this fragment revealed only one major cap site (Fig. 4b, lane 2). The amount of probe protected in this experiment was substantially less than that obtained with probes for *REP1* or *REP2* cap sites and probably reflects the relatively low abundance of *FLP* mRNA in [cir$^+$] yeast cells. The protected fragment corresponds to a cap site at position 5549, 21 bp upstream from the open reading frame. Thus, the 5′ termini of *REP2* and *FLP* transcripts lie at least 326 bp apart and quite close to the start of their respective reading frames.

Identification of polyadenylation sites for the major 2μm circle genes. Although several studies have recently addressed the nature of transcription termination or polyadenylation sites in *S. cerevisiae* (1, 16, 17, 37), there appears to be little consensus among these studies as to the specific sequences or structural features required for these events. We decided to determine empirically the locations of the 3′ ends of the major 2μm circle transcripts, anticipating that these data would provide additional material with which to assess 2μm circle transcription and its regulation, as well as to evaluate the general applicability of proposed polyadenylation-termination consensus sequences.

We identified the 3′ ends of 2μm circle transcripts by using an S1 nuclease protection protocol analogous to that used to determine the 5′ ends. In this case, though, we isolated fragments spanning the carboxy ends of the three genes and then used the DNA polymerase Klenow fragment and [α-^{32}P]deoxyribonucleoside triphosphates to label the 3′ ends of the probes. We then hybridized separated strands of the probes to poly(A)$^+$ RNA, and then we digested the residual single-strand nucleic acid with S1 nuclease. The size of each protected fragment was determined by electrophoresis on denaturing acrylamide gels.

To localize the 3′ end of *REP1* transcripts, we used a 912-bp *Hin*dIII fragment that spans an inverted repeat as well as the carboxy end of the *REP1* coding region. The size of the fragment protected by *REP1* RNA is shown in Fig. 5a (lane 5). One major fragment 184 bases long is evident, as are a number of smaller and larger fragments, which are present in reduced amounts. The size of the predominant fragment indicates the existence of a polyadenylation site at position 833, 58 bp downstream from a TAG codon that terminates the *REP1* open reading frame. The minor fragments may represent secondary polyadenylation sites at positions 827 and 859. When we used the opposite strand of this 3′-end-labeled *Hin*dIII fragment as probe, we observed protection of a small amount of a fragment more than 400 bases long (Fig. 5a, lane 3). This fragment most likely arose by protection conferred by *FLP* RNA, because the 3′ end of this *Hin*dIII fragment lies within the *FLP* coding region (Fig. 5c). The amount of this fragment, as judged by its intensity compared with that of the *REP1*-protected fragment, suggests that the steady-state level of *FLP* RNA is significantly less than that of *REP1* RNA in exponentially growing [cir$^+$] strains. This conclusion is consistent with our previous Northern analysis of 2μm circle transcripts (5).

We obtained a more precise localization of the *FLP* polyadenylation site by using a *Hin*dIII-*Xba*I fragment as probe (Fig. 5c). As in the previous experiment, we observed only one protected fragment (Fig. 5b, lane 2). This fragment,

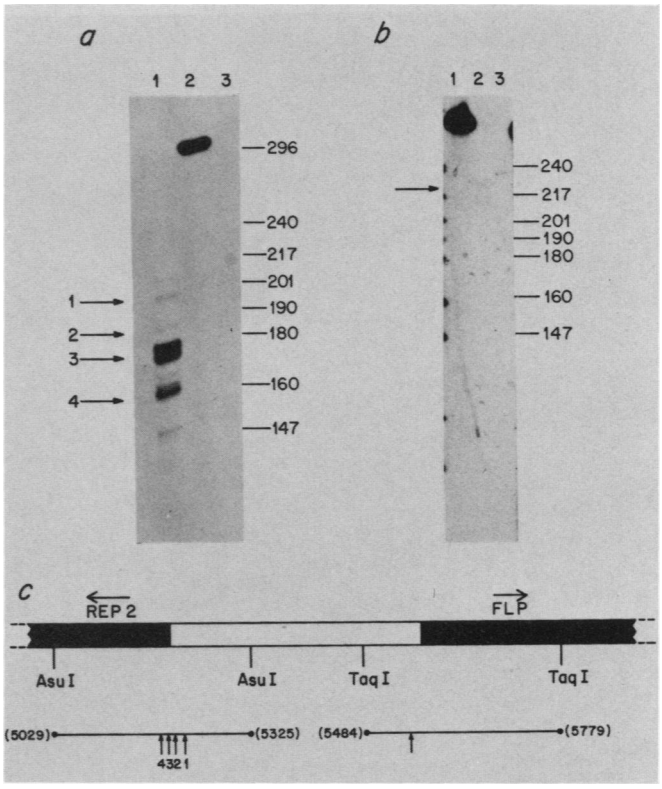

FIG. 4. S1 nuclease mapping of the 5′ ends of the *REP2* and *FLP* transcripts. S1 mapping was performed with the probes indicated in panel c as described in the legend to Fig. 2. Both separated strands of each DNA fragment were used as probes. Results are shown only for that strand for which specific protection was observed. (a) Identification of *REP2* cap sites by using an *Asu*I fragment (positions 5029 to 5325) as probe. Lanes: 1, probe hybridized to [cir$^+$] RNA and treated with 100 U of S1 nuclease; 2, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 3, probe hybridized to [cir^0] RNA, 100 U of S1 nuclease added. Numbers to the right of the figure designate the sizes in base pairs and positions of migration of pBR322 *Hpa*II fragments cofractionated on the same gel (data not shown). Arrows identify the major fragments protected from S1 nuclease digestion by *REP2* RNA. (b) Mapping of *FLP* RNA with a *Taq*I fragment (positions 5484 to 5779) as probe. Lanes: 1, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 2, probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added; 3, probe hybridized to [cir^0] RNA, 100 U of S1 nuclease added. Arrow points to the position of a faint band in lane 2 that resulted from protection of the probe by *FLP* RNA. (c) Location of probes used for 5′ mapping of *REP2* and *FLP* RNAs. Arrows indicate positions of the termini identified in this experiment.

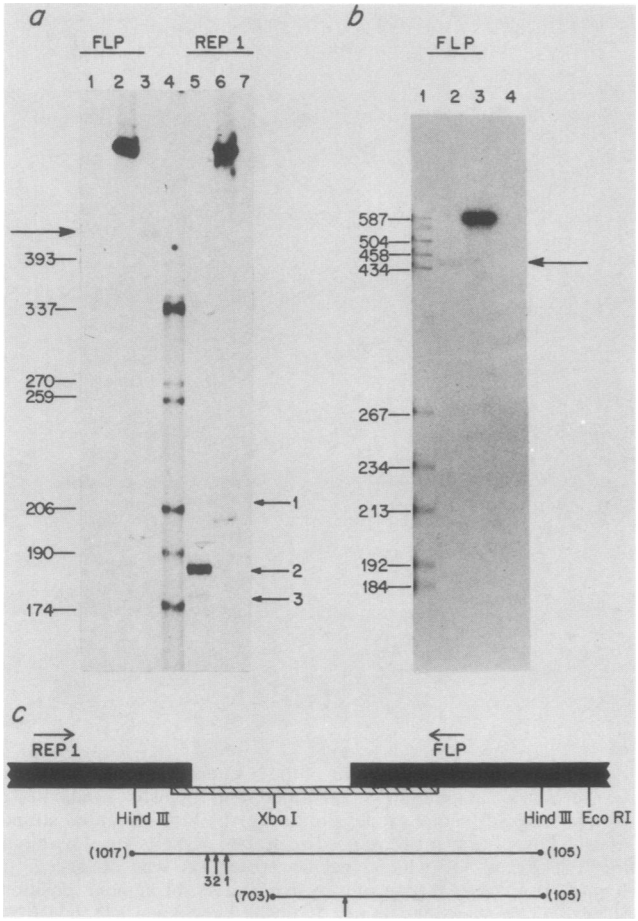

FIG. 5. Localization of the 3′ ends of the *REP1* and *FLP* transcripts. (a) S1 nuclease analysis with separated strands of a 3′-end-labeled *Hin*dIII fragment (positions 105 to 1017) as probe. Lanes: 1 to 3, results obtained by using the slower-migrating strand of the DNA probe; 5 to 7, results obtained by using the faster-migrating strand of the DNA probe; 1, probe hybridized to 25 μg of poly(A)$^+$ RNA from strain DCO4 [cir^0], 100 U of S1 nuclease added; 2, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 3, probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added; 4, pBR322 DNA digested with *Hha*I and 3′-end labeled; 5, bottom strand of probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added; 6, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 7, probe hybridized to [cir^0] RNA, 100 U of S1 nuclease added. Arrows on the right mark positions of fragments protected from nuclease digestion by *REP1* RNA. Arrow on the left marks the position of the fragment (lane 3) protected by *FLP* RNA. (b) Localization of the 3′ end of the *FLP* RNA by using an *Xba*I-*Hin*dIII fragment (positions 105 to 703) as probe. The two strands of the fragment were used separately as probes. Results are shown only for that strand for which specific protection was obtained. Lanes: 1, *Hin*fI digest of pBR322 DNA; 2, probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added; 3, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 4, probe hybridized to [cir^0] RNA, 100 U of S1 nuclease added. Arrow points to fragment (lane 2) protected from digestion by *FLP* RNA. (c) Location of probes from the 3′ ends of *REP1* and *FLP*. Hatched area represents one of the inverted repeats of the 2μm circle. Arrows show the position of the polyadenylation sites identified in panels a and b.

440 bases long, corresponds to a polyadenylation site at position 545, 24 bp downstream from the nonsense codon at the end of the *FLP* coding region.

Identification of the polyadenylation site for *REP2* was accomplished by using as probe an *Xba*I-to-*Asu*I fragment that extends from position 3945 to 4387 (Fig. 6b). The terminator of the *REP2* coding region lies at position 4285. S1 nuclease analysis (Fig. 6a, lane 3) yielded two major protected fragments 278 and 282 bp long. These correspond to polyadenylation sites at positions 4105 and 4109, which are in the inverted repeat about 178 bp downstream from the translation termination codon for the gene. This distance is unusually large when compared with those of the 3′ untranslated regions for *REP1* (58 bp) and *FLP* (24 bp). However, the *REP2* polyadenylation site lies at exactly the same sequence as that for the *FLP* gene but in the opposite inverted repeat. It thus appears that sequence specificity is

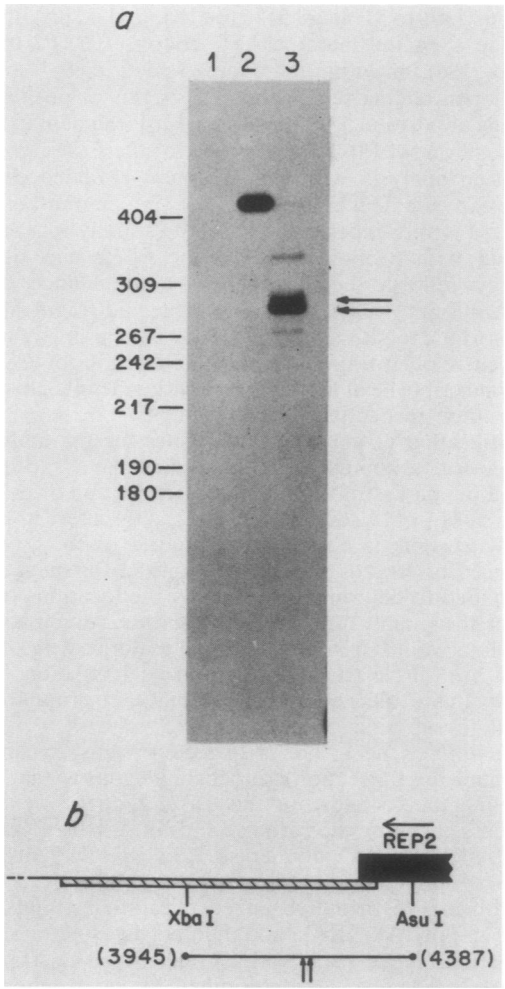

FIG. 6. Localization of the 3′ end of the *REP2* transcript. (a) S1 nuclease mapping by using a 3′-end-labeled *Xba*I-*Asu*I fragment as probe. The two strands of the fragment were used separately as probes. Results are shown only for that strand for which specific protection was obtained. Lanes: 1, probe hybridized to [cir^0] RNA, 100 U of S1 nuclease added; 2, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 3, probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added. Numbers refer to the positions of fragments of pBR322 DNA digested with *Hae*III or *Hpa*II, end labeled, and fractionated on the same gel (data not shown). Arrows point to the two major fragments protected from S1 nuclease by *REP2* RNA. (b) Location of the probe from the 3′ end of the *REP2* gene. Hatched area represents one of the inverted repeats of the plasmid. Arrows mark the position of the polyadenylation sites mapped in this experiment.

of prime importance in determining the site of polyadenylation in the plasmid.

Transcription of coding region D. Coding region D is a 543-bp open reading frame that extends from an ATG codon at position 2271 to a TAA codon at position 2814. No phenotype has been ascribed to disruptions of this coding region. Nonetheless, we were interested in determining the extent to which this coding region constitutes a functional gene. To this end, we examined the question of whether a specific transcript is derived from this region in vivo. From the size of the coding region, we anticipated a transcript of at least 600 bases, the exact size depending on the length of the nontranslated segments carried on the transcript and the length of the poly(A)$^+$ tail. Our previous Northern analysis of 2μm circle transcripts identified two 700-bp RNA species in [cir$^+$] cells that were homologous to the large unique region (5). However, the precise genomic regions homologous to these transcripts were not unambiguously established.

To learn whether there is a transcript functionally encompassing the D coding region, we used S1 nuclease analysis to determine the 5′ and 3′ ends of RNA species spanning D. First, to examine whether there are transcripts whose 5′ ends correspond to the 5′ end of the D coding region, we used as a probe for S1 analysis a 5′-end-labeled fragment extending from a *Pvu*II site at position 2133 to an *Hin*dIII site at position 2331 (Fig. 7c). The results of this analysis (Fig. 7a, lane 2) show that RNA from a [cir$^+$] strain confers protection to a portion of this fragment. On the basis of the sizes of the protected regions, we concluded that these cells contain transcripts whose 5′ ends correspond to positions between 2251 and 2257, or sites 14 to 20 bp upstream from the initial ATG of the D coding region. By using a 3′-end-labeled *Eco*RI-to-*Hpa*I fragment as probe in S1 analysis of the same RNA, we also identified transcripts whose 3′ ends lie near the carboxy end of D. The results of this experiment (Fig. 7b, lane 3) indicate the existence of polyadenylation sites at positions 2841 and 2860. These two sets of experiments taken together document that in [cir$^+$] strains there are transcripts that extend from 17 bp upstream of the initial ATG of the D coding region to either 27 or 46 bp downstream of the codon ending the coding region. Thus, these data in conjunction with our previous Northern analysis, which demonstrated the existence of a 600- to 700-bp transcript within the large unique region, suggest that the D coding region in fact constitutes a functional, transcribed gene.

In addition to conferring protection to a distinct subdomain of the 5′-end-labeled *Pvu*II-to-*Hin*dIII fragment used to map the 5′ end of the D transcript, RNA from a [cir$^+$] strain also conferred protection to a significant proportion of the full-length probe. This protection most likely resulted from hybridization of the probe to the 1,620-base 2μm circle transcript. That is, our previous Northern analysis (5) indicated that this transcript should completely span the probe. In addition, the absence of any protection of the full-length fragment after hybridization with RNA from a [cir^0] strain (Fig. 7a, lane 3) argues that residual probe present (lane 2) was not an artifact resulting from incomplete digestion. Because there was no protection of full-length probe with the 3′-end-labeled fragment (Fig. 7b, lane 3) and no protected fragments other than those corresponding to the 3′ end of the D coding region, we assume that the 1,620-base transcript and the transcript for D are coterminal.

Deletion analysis of a polyadenylation site in 2μm circle. From our previous Northern analysis of 2μm circle transcription and from our data obtained by S1 nuclease analysis, we have found that the 1,620-base and D transcripts have 3′ ends between the *Pst*I and *Hpa*I sites (Fig. 7 and 8). In our previous analysis of the physical limits of the *REP3* locus, we constructed plasmids containing small deletions centered around this *Hpa*I site (20). We were, accordingly, interested in examining the transcription pattern of plasmids containing these deletions as a means of evaluating the sequence requirements of polyadenylation-termination in 2μm circle.

We isolated poly(A)$^+$ RNA from a [cir^0] strain either harboring plasmid CV20, which consists of the entire 2μm

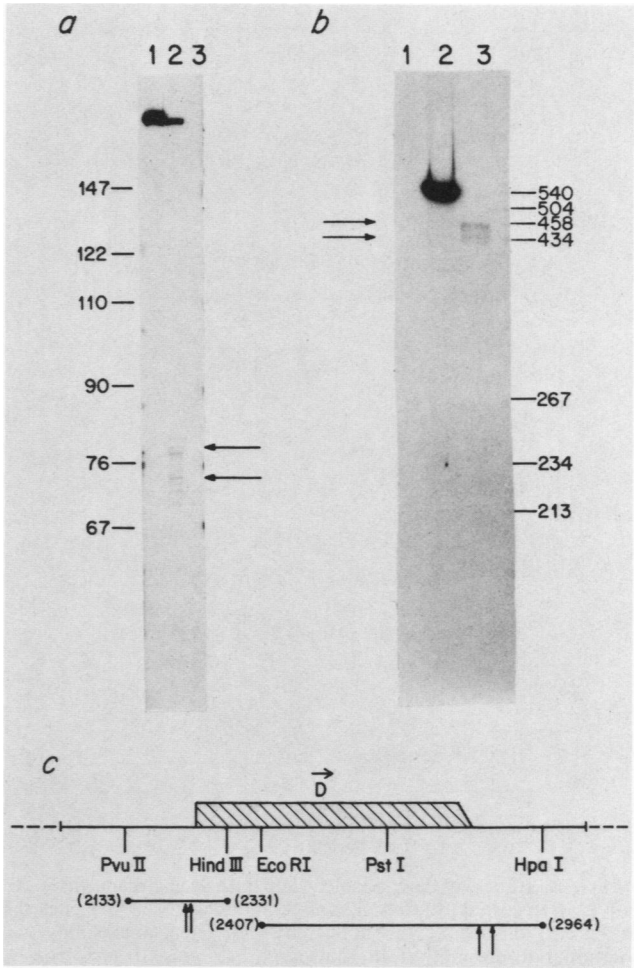

FIG. 7. Localization of the 5′ and 3′ termini of the transcript of coding region D. Mapping was performed as described in the legends to Fig. 2 and 5 and in Materials and Methods. Both separated strands of each DNA fragment were used as probes. Results are shown only for that strand for which specific protection was observed. (a) Identification of the 5′ end by using a *Pvu*II-*Hin*dIII fragment (positions 2133 to 2331) as probe. Lanes: 1, probe hybridized to [cir$^+$] RNA, no S1 nuclease added; 2, probe hybridized to [cir$^+$] RNA, 100 U of S1 nuclease added; 3, probe hybridized to [cir^0] RNA, 100 U of S1 nuclease added. Numbers refer to positions of fragments of pBR322 DNA digested with *Hpa*II and run on the same gel (data not shown). Arrows mark two of the major fragments protected from digestion by the D transcript. (b) Identification of the 3′ end by using an *Eco*RI-*Hpa*I fragment (positions 2407 to 2964) as probe. Lanes: 1, [cir^0] RNA, 100 U of S1 nuclease; 2, [cir$^+$] RNA, no S1 nuclease; 3, [cir$^+$] RNA, 100 U of S1 nuclease. Numbers are from an *Hae*III digest of pBR322 DNA run on the same gel. Arrows mark the two major fragments protected by RNA from coding region D. (c) Location of probes used and termini mapped in this experiment.

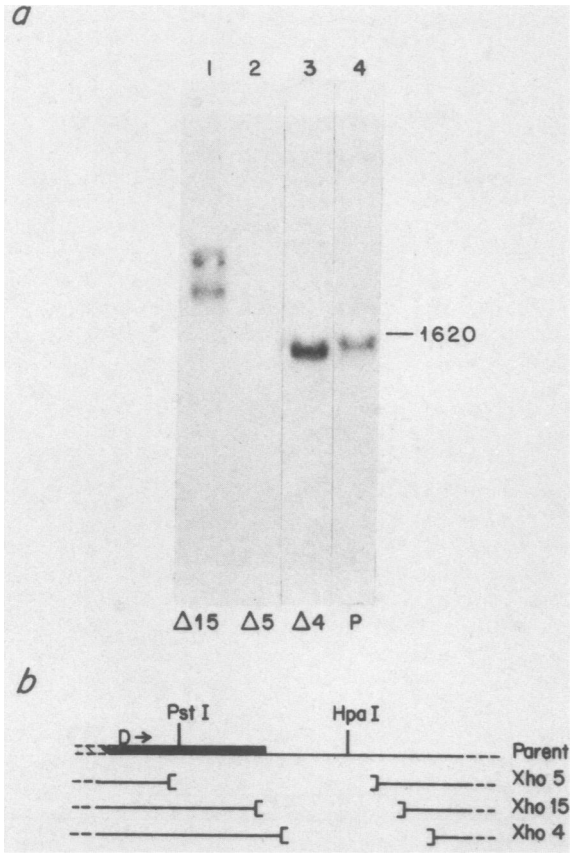

FIG. 8. Effects of deletions near *REP3* on 2μm circle transcripts. Poly(A)$^+$ RNA (20 μg) from a strain containing CV20 (designated P or Parent) or one of the indicated Xho deletion plasmids was fractionated on methyl mercury agarose gels, transferred to diazotized paper, and probed as described in Materials and Methods. Lanes 1 to 4 were hybridized with a single-strand DNA fragment specific for transcripts extending from the left into the region shown in panel b. Concurrently, fractionated standards (not shown) allowed determination of the sizes of the transcripts. (b) Deletion endpoints of Xho plasmids used in this study. The regions deleted in plasmids Xho4, Xho5, and Xho15 are indicated below the diagram of that portion of the large unique region of 2μm circle spanning the *Hpa*I site. The 3′ portion of the D coding region lies at the left, and the *REP3* locus (not labeled) lies at the right. The *REP1* gene, from which the probe used in panel a was obtained, lies outside the figure to the left.

circle genome with pBR322 plus *LEU2* cloned into the small unique region, or harboring members of our Xho series of plasmid CV20 derivatives. These plasmids are identical to plasmid CV20 except for deletions around the *Hpa*I site whose endpoints are indicated in Fig. 8b. We fractionated samples of RNA from these strains on denaturing agarose gels and transferred the fractionated RNA to diazobenzyloxymethyl-paper. We then hybridized the immobilized RNA with one strand of a DNA fragment derived from the *REP1* coding region, that is, from that region of the 2μm circle genome lying immediately to the left of the sequence diagrammed in Fig. 8b. The strand used hybridized only to those transcripts extending from the left into the region shown. Thus, only transcripts from the 1,620-base transcription unit were visualized by this procedure.

The effects of the Xho deletions on termination of the 1,620-base transcript are evident in Fig. 8a. Normal levels of this transcript are seen in strains harboring plasmids CV20 or Xho4. However, this transcript is absent in strains harboring plasmid Xho5 or Xho15. In addition, in *S. cerevisiae* carrying plasmid Xho15, the 1,620-base transcript is replaced by several substantially longer transcripts. These data suggest that the signals for polyadenylation or termination of the 1,620-base transcript are deleted in plasmids Xho15 and Xho5 but retained in plasmid Xho4. In addition, it appears that the resulting readthrough transcripts from plasmid Xho15 are reasonably stable and terminate at specific sites downstream, whereas those from plasmid Xho5 are unstable.

DISCUSSION

By S1 nuclease and primer extension mapping procedures, we defined precisely the cap and polyadenylation sites of transcripts from the four genes in the yeast 2μm circle plasmid. Their positions are given in Table 1, along with the calculated transcript length for each gene. Assuming a 50-bp poly(A)$^+$ tail, the predicted lengths for the four transcripts are somewhat shorter than, but in the same relative order as, those we previously determined by Northern analysis. In this study we also identified the polyadenylation sites of two other plasmid transcripts of unknown function and identified by deletion analysis the sequences required for effecting this process. Our results allow us to address several issues concerning transcriptional expression of the plasmid, as well as the nature of transcriptional signals in yeast cells in general. In addition, these results provide a framework for continued studies on the regulation of gene expression of the 2μm circle.

2μm circle genes are transcribed independently. We previously proposed that the transcription of genes in the small unique region, as well as that of *REP1*, is initiated at a single site in the large unique region of the plasmid—at the promoter for the 1,950-base transcript (5). The observed gene-sized transcripts for the three large open reading frames, as well as the 1,950-base RNA, would thus represent processed products of two larger genome-sized RNAs. The two large transcripts were proposed to be derived from the two forms of the plasmid that arise from *FLP*-mediated intramolecular recombination. However, the results presented here are more readily accommodated assuming that the four plasmid genes are independently transcribed and that the 5′ and 3′ ends identified for plasmid transcripts represent authentic initiation and termination sites.

A central tenet of the previous model is a precursor-

TABLE 1. Cap and polyadenylation sites for the major 2μm circle genes[a]

Gene	Cap site	Poly-adenylation site	Transcript size (bp)
REP1	2008, 2035	833	1,200
REP2	5201–5198, 5186–5184	4107	1.090
FLP	5549	545	1,305
D	2251, 2257	2841, 2860	600

[a] Positions of the major cap and polyadenylation sites for each of the genes were obtained as described in the text and are identified by the numbering system of Hartley and Donelson (13) for the A form of 2μm circle. We estimate that our assignment of the positions of cap and polyadenylation sites is accurate to within 3 bp. Transcript size represents the approximate distance in base pairs between the major cap and polyadenylation site for each gene. No allowance was made for polyadenylation, so the actual in vivo transcript is somewhat larger than the value given.

product relation between the 1,950- and 1,325-base transcripts. However, we were unable to detect by S1 nuclease mapping an RNA species with a 3′ terminus near the 5′ end of *REP1*. Such a species would be expected as a by-product of the nucleolytic cleavage of the 1,950-base transcript that would yield the 1,325-base RNA. It is possible, though, that this species would be rapidly degraded and thus not detectable as a stable RNA. A more compelling indictment of the proposed precursor-product relationship, however, is the presence of normal levels of the 1,325-base transcript in strains harboring plasmid Xho4, although the deletion in this plasmid completely abolishes any accumulation, presumably by abolishing synthesis, of the 1,950-base transcript (21). Therefore, it is unlikely that the *REP1* mRNA is derived by processing of the longer transcript.

Consistent with the concept of independent transcription of the major plasmid genes is the existence of promoterlike sequences appropriately positioned relative to the mRNA cap sites of the three genes. The sequence TATAAA, which is identical to the consensus yeast TATA box (35), lies 66 bp upstream from the cap site of the longest *REP1* mRNA. This distance is consistent with that found for other yeast promoter regions. Similarly, the sequence TATAAT lies 35 bp upstream of the cap site of the longest *REP2* mRNA, and the sequence TAAATA lies 32 bp upstream of the single *FLP* gene cap site. It is possible that the deviation of the *FLP* TATA-like sequence from the consensus sequence is in some way related to the relatively low level of expression of the gene. Other components of 2μm circle promoters are less readily apparent by simple visual inspection of sequence, as is true of any yeast promoter, and so their identification awaits further analysis.

Translation of *REP1*. The positions of the multiple *REP1* mRNA cap sites pose some ambiguity regarding the means by which these transcripts are translated. The cap sites for predominant *REP1* transcripts lie either within the *REP1* coding region or upstream from an AUG codon that is followed almost immediately by a UGA termination codon (Fig. 3). Only a few transcripts have cap sites that lie between this upstream AUG and the initial AUG of the coding region. Strict application of the scanning hypothesis for eucaryotic translation (23–26) would then suggest either that most of the *REP1* transcripts are not translationally competent for the synthesis of *REP1* protein or that significant suppression of the intervening UGA nonsense codon is required to allow readthrough of translation initiated at the upstream AUG. Although there is evidence for the existence of UGA suppressor activity in tRNA isolated from wild-type yeast cells and for UGA suppressor activity in higher cells (10, 14), the level of in vivo UGA suppression in wild-type strains is quite low (15). In addition, recent analysis of the sequence of the amino end of the *REP1* protein indicates that most, if not all, of the protein arises by initiation of translation at the initial AUG of the coding region (L. Wu and J. R. Broach, unpublished observations).

Recent reexamination of the scanning hypothesis provides an alternative view of the mechanism by which *REP1* transcripts are translated (27, 28). That is, it appears that reinitiation of translation at a second AUG of a eucaryotic mRNA can occur in some instances, provided that the first open reading frame is short. In fact, for the *src* gene of Rous sarcoma virus, such reinitiation appears to be the only means by which translation of the coding region is initiated on *src* gene mRNA (19). This, then, provides a plausible mechanism for *REP1* translation: reinitiation of translation at the second AUG would allow synthesis of the *REP1* protein from those transcripts with cap sites upstream from the prior AUG. The function, if any, of the short upstream coding region is certainly unknown.

We observed considerable heterogeneity in the cap sites of transcripts from most of the 2μm circle genes. Similar heterogeneity in 5′ termini has been described for a number of yeast genes (7, 33–35). In some cases, changes in the pattern of 5′ termini have been correlated with changes in gene expression (7, 34). For *REP1*, differential regulation of the site of transcription initiation would affect which of the two AUG codons was used to initiate translation, with the attendant consequence of translating or not translating the short upstream reading frame, or, if the analogy to *src* gene translation is true, of translating or not translating the *REP1* gene itself.

Consensus sequences for polyadenylation in 2μm circle. The specific signals responsible for polyadenylation and transcription termination in *S. cerevisiae* are not well defined. By comparison of the sequences at the 3′ end of a number of yeast genes, Bennetzen and Hall noted a segment, lying 25 to 40 bases upstream of the site of polyadenylation of each gene, whose sequence exhibited a recurrent theme (1). They suggested that the derived consensus sequence of these segments, TAAATAAA(or G), or some variations of this sequence, provides a signal for polyadenylation in yeast transcripts. In a separate study, Zaret and Sherman identified a deletion mutation at the 3′ end of the *CYC1* gene that abolishes the normal polyadenylation site of the *CYC1* transcript and results in accumulation of longer readthrough transcripts of various lengths (37). They compared the sequences deleted by this mutation with those near the polyadenylation sites of other yeast genes and proposed a tripartite signal for polyadenylation whose consensus sequence is TAG. . .TATG. . .TTT or TAG. . .TATGT . . .TTT. Finally, on the basis of the positions of the 3′ termini of transcripts from various constructions of a *Drosophila melanogaster* gene expressed in yeast cells, Henikoff et al. proposed that the sequence TTTTTATA 50 to 90 bp upstream from the 3′ end constitutes part of the signal for polyadenylation (16, 17).

In none of the cases examined has polyadenylation been distinguished from termination. That is, in none of these systems has transcription been examined by pulse-labeling or by nuclear runoff, to determine whether transcription continues beyond the site of polyadenylation. Thus, it is not now possible to assess whether the sequences identified can be considered to signal transcription termination with attendant polyadenylation or only endonucleolytic cleavage and polyadenylation of an elongating transcript. The latter situation prevails in many higher eucaryotic transcription units (18).

The sequences surrounding the polyadenylation sites of the four open coding regions of the 2μm circle plasmid are shown in Table 2. There are two polyadenylation sites at the 3′ end of the D coding region. It is not clear whether the two sites arise as a result of two separate signals or as a result of the ambiguous execution of a single weak signal. For completeness, the sequence spanning both sites is displayed. As is evident from the highlighted portions of these sequences, all of them encompass a region resembling the tripartite polyadenylation signal proposed by Zaret and Sherman (37). Sequences similar to the consensus elements proposed by Bennetzen and Hall (1) or Henikoff et al. (17) are not readily apparent.

The effects of deletions on the polyadenylation sites at the 3′ end of the D coding region provide additional significance

TABLE 2. Putative polyadenylation signals at the 3′ end of 2 μm circle genes[a]

Gene	Signal
REP1	.....TAGGTTATATAGGGATA**TAG**CACAGAGA**TATAT**AGCAAAGAGATACTTTTGAGCAA .....
REP2	..TGATAG .. (143bp) .. CCTACATAAA**TAG**ACGCATA**TAAGT**ACGCATTTAAGCATAAACACGC.....
FLP	.....CCTACATAAA**TAG**ACGCATA**TAAGT**ACGCATTTAAGCATAAACACAC .....
	← 15 ← 4
D	.....**TAGT**ATATTTGAACCTGTATAATAATATA**TAGT**CTAGCGCTTTACGGAAGACAATGTATGTATTTCGG .....
D(b)	.....TAGTC**TAG**CGCTTTACGGAAGACAATG**TATGT**ATTTCGGT .....
	TAA
Consensus	.....TAG .....(1–140 bp).....**TAG** .....**TA(T)GT** ..... TTT .....
	TGA

[a] The sequences immediately upstream of the polyadenylation site for each of the 2μm circle genes are compared with the consensus polyadenylation signal identified by Zaret and Sherman (37). The translation termination codon for each gene is underlined, and the polyadenylation site is marked by an exclamation point. The specific sequences exhibiting homology to the consensus sequence are shown in boldfaced type. The left-hand endpoints of deletions in plasmids Xho15 and Xho4 are indicated above the sequence of the 3′ end of the D coding region. The sequence shown for *REP1* extends from position 889 to 834, the sequence for *REP2* extends from position 4152 to 4106 (excluding the initial hexanucleotide), the sequence for *FLP* extends from position 501 to 545, the sequence for D extends from position 2795 to 2863, and the sequence for D(b) extends from position 2824 to 2864.

to the sequences identified by Zaret and Sherman (37). The results presented in Fig. 8 show that sequences lying between the endpoints of deletions Xho4 and Xho15, the precise locations of which are indicated in Table 2, are necessary for effecting efficient maturation of the 3′ end of the 1,620-base transcript. That is, normal-length 1,620-base transcripts are present in cells harboring plasmid Xho4, whereas the 1,620-base transcript is absent, and longer, readthrough transcripts are present in cells containing plasmid Xho15. Significantly, the region lying between the left-hand endpoints of the deletions in plasmids Xho4 and Xho15 encompasses the tripartite signal. It is noteworthy that all of the readthrough transcripts are poly(A)$^+$. There are no transcripts from the 1,620-base transcription unit in total RNA other than those seen in poly(A)-selected RNA (data not shown). Also noteworthy is the absence of readthrough transcription to the downstream polyadenylation sites in cells carrying plasmid Xho4. This is true even though the deletion in plasmid Xho4 removed the second polyadenylation site at the 3′ end of D. This suggests that the complete signal for polyadenylation lies upstream from the actual site of polyadenylation, as is true for higher eucaryotes and as has been suggested for yeasts (8, 17).

Deletion of sequences at the 3′ end of coding region D appears to affect the stability of D transcripts as well as the maturation of their 3′ ends. No transcripts from the 1,620-base transcription unit, either of the normal size or of a size expected for readthrough transcription, are seen in cells containing plasmid Xho5 (Fig. 8). This is true even though approximately normal levels of transcription occur from the opposite strand. Thus, either the particular junction sequence generated by the deletion is inimical to the stable maintenance of transcripts spanning it or the sequences deleted by the mutation are required for stable maintenance.

The D coding region. Evidence presented in this paper indicates that coding region D is transcribed into a 600-bp poly(A)$^+$ RNA whose 5′ end lies near the start of the open coding region. Thus, it appears that this region constitutes a functional gene. From what is known of other 2μm circle genes, one would anticipate that such a gene would be likely to play some role in plasmid maintenance. As of yet, however, no obvious deficiency in plasmid maintenance has been observed as a consequence of substantial disruptions of this coding region. However, it should be noted that previous genetic characterizations of the function of the D coding region have been conducted by using hybrid plasmids. Such plasmids generally consist of large insertions into the 2μm circle genome of bacterial and yeast chromosomal sequences that provide functions necessary for propagation in *E. coli* and selection in *S. cerevisiae*. However, these hybrid plasmids are significantly less stable than authentic 2μm circle. It is not clear whether the reduction in stability is the result of some form of *cis*-active inhibition induced merely by the presence of the inserted sequences or whether the plasmid is so tightly organized that there are essentially no neutral sites into which to insert foreign DNA. In either case, though, if the D coding region encodes a product required in some subtle fashion for enhanced plasmid maintenance, the phenotypic consequence of inactivating the gene might well be obscured by the more substantial perturbations caused by the presence of sequences used to construct the hybrid plasmids. Such a situation appears to be the case with the *FLP* gene. Recent evidence suggests that the product of this gene is essential in effecting plasmid amplification (Murray and Szostak, personal communication). Nonetheless, inactivation of *FLP* does not yield any apparent diminution in the

persistence of most hybrid plasmids examined (20, 31). Thus, it will be of interest to reexamine the consequences of inactivating the D region in plasmids less extensively perturbed than those used previously and to perform these experiments by using assays that are more sensitive to subtle alterations in plasmid properties.

Regulation of 2μm circle transcription. The steady-state levels of 2μm circle transcripts we observed in this study, as assessed by the degree of protection afforded by total [cir$^+$] yeast RNA against S1 nuclease digestion of appropriate probes, corresponded well with the levels we previously observed with Northern analysis. In both cases, however, we were examining the levels only in exponentially growing yeast cells. We anticipate that these levels fluctuate in different conditions, that is, that the expression of 2μm circle genes is regulated.

The 2μm circle plasmid encodes a number of activities that ensure plasmid maintenance during mitotic growth and throughout meiosis. These activities include plasmid partitioning, promoted by the products of the *REP1* and *REP2* genes, and plasmid amplification, apparently induced by the product of the *FLP* gene (20–22, 36; Murray and Szostak, personal communication). It also appears that plasmid copy levels are actively maintained, because 2μm circle exhibits plasmid incompatibility. However, this copy control system is not so rigid as that observed with most bacterial plasmids; there appears to be some clonal variation in copy number. It is possible, as suggested by Futcher and Cox (9), that plasmid copy level is established passively as an equilibrium between opposing influences on plasmid copy number. That is, a constant tendency toward increased copy levels, as a result of some activity promoting hyperreplication, for example, could be counterbalanced by other activities promoting decreased copy levels. Such an activity could be selection against cells with high plasmid copy numbers, for example. On the other hand, recent observations by Murray and Szostak (personal communication) suggest that *FLP*-promoted amplification can induce a rapid and substantial increase in copy number. Thus, in this case, modulation of at least *FLP* gene expression would be required to avoid excessive plasmid accummulation and to maintain reasonable copy levels. The low levels of *FLP* expression we see in exponentially growing yeast cells may reflect repression of *FLP* expression during steady-state growth conditions.

Several recent observations suggest that 2μm circle transcription is, in fact, regulated. Veit and Fangman have noted specific alterations in plasmid chromatin structure near the 5′ ends of various 2μm genes as a function of the presence or absence of various 2μm circle gene products (B. E. Veit and W. L. Fangman, personal communication). They suggest that these chromatin alterations may reflect specific regulatory interactions. Similarly, we have observed that the *leu2-d* gene of plasmid pJDB219, which consists of the *LEU2* coding region fused to a 2μm circle promoter, is transcribed at a significantly lower rate in [cir$^+$] strains than in [cir^0] strains (T. Som, A. Sutton, and J. R. Broach, unpublished observations). This suggests that the 2μm circle promoter responsible for expression of *leu2-d* is negatively regulated by one or more products of the plasmid itself. Consistent with this hypothesis is our observation that deletion of sequences near this promoter yields substantially enhanced production of RNA initiated at this site (21). Thus, we anticipate that 2μm circle transcription will exhibit substantial regulation and that this regulation will be intimately involved in maintaining plasmid stasis during normal growth and ensuring plasmid persistence after significant perturbations of plasmid levels from the norm. The transcription mapping presented here should be of value in pursuing these studies.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM34596 from the National Institutes of Health. J.R.B. is an Established Investigator of the American Heart Association, and A.S. is a postdoctoral fellow of the National Institutes of Health.

LITERATURE CITED

1. **Bennetzen, J. L., and B. D. Hall.** 1982. The primary structure of the *Saccharomyces cerevisiae* gene for alcohol dehydrogenase I. J. Biol. Chem. **257**:3018–3025.
2. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. USA **75**:1274–1278.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
4. **Broach, J. R.** 1981. The yeast plasmid 2 micron circle, p. 445–470. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*, vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Broach, J. R., J. F. Atkins, C. McGill, and L. Chow.** 1979. Identification and mapping of the transcriptional and translational products of the yeast plasmid, 2μ circle. Cell **16**:827–839.
6. **Broach, J. R., and J. B. Hicks.** 1980. Replication and recombination functions associated with the yeast plasmid 2μ circle. Cell **21**:501–508.
7. **Faye, G., D. W. Leung, K. Tatchell, B. D. Hall, and M. Smith.** 1981. Deletion mapping of sequences essential for *in vivo* transcription of the iso-1-cytochrome *c* gene. Proc. Natl. Acad. Sci. USA **78**:2258–2262.
8. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5′-AAUAAA-3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
9. **Futcher, A. B., and B. S. Cox.** 1984. Copy number and the stability of 2-μm circle-based artificial plasmids of *Saccharomyces cerevisiae*. J. Bacteriol. **157**:283–290.
10. **Geller, A. I., and A. Rich.** 1980. A UGA termination suppressor $tRNA^{Trp}$ active in rabbit reticulocytes. Nature (London) **283**:41–46.
11. **Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. Lebowitz, and S. M. Weissman.** 1978. Heterogeneity and 5′-terminal structures of the late RNAs of simian virus 40. J. Mol. Biol. **126**:813–846.
12. **Guo, L.-H., and R. Wu.** 1983. Exonuclease III: use for DNA sequence analysis and in specific deletions of nucleotides. Methods Enzymol. **100**:60–96.
13. **Hartley, J. L., and J. E. Donelson.** 1980. Nucleotide sequence of the yeast plasmid. Nature (London) **286**:860–865.
14. **Hatfield, D., A. Diamond, and B. Dudock.** 1982. Opal suppressor serine tRNAs from bovine liver form phosphoseryl-tRNA. Proc. Natl. Acad. Sci. USA **79**:6215–6219.
15. **Hawthorne, D.** 1981. UGA suppressors in yeast, p. 291–301. *In* D. von Wettstein, J. Friis, M. Kielland-Brandt, and A. Stenderup (ed.), Molecular genetics of yeast. Alfred Benzon symposium 16. Munksgaard, Copenhagen.
16. **Henikoff, S., and E. H. Cohen.** 1984. Sequences responsible for transcription termination on a gene segment in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:1515–1520.
17. **Henikoff, S., J. D. Kelly, and E. H. Cohen.** 1983. Transcription terminates in yeast distal to a control sequence. Cell **33**:607–614.
18. **Hofer, E., and J. E. Darnell.** 1981. The primary transcription unit of the mouse β-major globin gene. Cell **23**:585–593.
19. **Hughes, S., K. Mellstrom, E. Kosik, F. Tamanoi, and J. Brugge.** 1984. Mutation of a termination codon affects *src* initiation. Mol. Cell. Biol. **4**:1738–1746.
20. **Jayaram, M., Y. Y. Li, and J. R. Broach.** 1983. The yeast plasmid 2μ circle encodes components required for its high copy propagation. Cell **34**:95–104.
21. **Jayaram, M., A. Sutton, and J. R. Broach.** 1985. Properties of

REP3: a *cis*-acting locus required for stable propagation of the *Sacharomyces cerevisiae* plasmid 2μm circle. Mol. Cell. Biol. **5**:2466–2475.
22. **Kikuchi, Y.** 1983. Yeast plasmid requires a *cis*-acting locus and two plasmid proteins for its stable maintenance. Cell **35**: 487–493.
23. **Kozak, M.** 1980. Evaluation of the scanning model for initiation of protein synthesis in eucaryotes. Cell **22**:7–8.
24. **Kozak, M.** 1981. Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. Curr. Top. Microbiol. Immunol. **93**:81–123.
25. **Kozak, M.** 1983. Translation of insulin-related polypeptides from messenger RNAs with tandemly reiterated copies of the ribosome binding site. Cell **34**:971–978.
26. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
27. **Kozak, M.** 1984. Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. Nucleic Acids Res. **12**:3873–3893.
28. **Liu, C.-C., C. C. Simonsen, and A. D. Levinson.** 1984. Initiation of translation at internal AUG codons in mammalian cells. Nature (London) **309**:82–85.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
31. **McLeod, M., E. Volkert, and J. R. Broach.** 1984. Components of the site-specific recombination system encoded by the yeast plasmid 2 micron circle. Cold Spring Harbor Symp. Quant. Biol. **49**:779–787.
32. **Murray, A. W., and J. W. Szostak.** 1983. Pedigree analysis of plasmid segregation in yeast. Cell **34**:961–970.
33. **Nasmyth, K. A., K. Tatchell, B. D. Hall, C. Astell, and M. Smith.** 1981. Physical analysis of mating-type loci in *Saccharomyces cerevisiae*. Cold Spring Harbor Symp. Quant. Biol. **45**:961–980.
34. **Rose, M., and D. Botstein.** 1983. Structure and function of the yeast URA3 gene. J. Mol. Biol. **170**:883–904.
35. **Sentenac, A., and B. D. Hall.** 1982. Yeast nuclear RNA polymerases and their role in transcription, p. 561–606. In J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*, vol. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Sigurdson, D. C., M. E. Gaarder, and D. M. Livingston.** 1981. Characterization of the transmission during cytoductant formation of the 2 micron DNA plasmid from *Saccharomyces cerevisiae*. Mol. Gen. Genet. **183**:59–65.
37. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28**:563–573.