Different sequence elements are required for function of the cauliflower mosaic virus polyadenylation site in Saccharomyces cerevisiae compared with in plants.

S Irniger, H Sanfaçon, C M Egli and G H Braus
Mol. Cell. Biol. **1992, 12(5):2322. DOI: 10.1128/MCB.12.5.2322.**

MOLECULAR AND CELLULAR BIOLOGY, May 1992, p. 2322–2330
0270-7306/92/052322-09$02.00/0

Vol. 12, No. 5

Different Sequence Elements Are Required for Function of the Cauliflower Mosaic Virus Polyadenylation Site in *Saccharomyces cerevisiae* Compared with in Plants

STEFAN IRNIGER,[1] HÉLÈNE SANFAÇON,[2]† CHRISTOPH M. EGLI,[1] AND GERHARD H. BRAUS[1]*

Institute of Microbiology, Swiss Federal Institute of Technology, Schmelzbergstrasse 7, CH-8092 Zürich,[1] and Friedrich Miescher Institute, CH-4002 Basel,[2] Switzerland

Received 26 November 1991/Accepted 3 February 1992

We show that the polyadenylation site derived from the plant cauliflower mosaic virus (CaMV) is specifically functional in the yeast *Saccharomyces cerevisiae*. The mRNA 3′ endpoints were mapped at the same position in yeast cells as in plants, and the CaMV polyadenylation site was recognized in an orientation-dependent manner. Mutational analysis of the CaMV 3′-end-formation signal revealed that multiple elements are essential for proper activity in yeast cells, including two upstream elements that are situated more than 100 and 43 to 51 nucleotides upstream of the poly(A) addition site and the sequences at or near the poly(A) addition site. A comparison of the sequence elements that are essential for proper function of the CaMV signal in yeast cells and plants showed that both organisms require a distal and a proximal upstream element but that these sequence elements are not identical in yeast cells and plants. The key element for functioning of the CaMV signal in yeast cells is the sequence TAGTATGTA, which is similar to a sequence previously proposed to act in yeast cells as a bipartite signal, namely, TAG...TATGTA. Deletion of this sequence in the CaMV polyadenylation signal abolished 3′-end formation in yeast cells, and a single point mutation in this motif reduced the activity of the CaMV signal to below 15%. These results indicate that the bipartite sequence element acts as a signal for 3′-end formation in yeast cells but only together with other *cis*-acting elements.

The generation of mRNA 3′ ends is an important step in gene expression. In eukaryotes, this process generally involves an endonucleolytic cleavage and the addition of a poly(A) tail. The identification of the sequences responsible for mRNA 3′-end formation in various eukaryotic organisms has revealed that mammals contain apparently simpler poly(A) signals than plants, retroviruses, and yeasts (for a review, see reference 33). Mammalian poly(A) signals are characterized by two *cis*-acting elements: the conserved sequence AATAAA, which is present 10 to 30 nucleotides upstream of the cleavage site (34), and a more-variant T-rich or GT-rich element immediately downstream of the cleavage site (8, 16, 26). Both elements are essential for proper cleavage and polyadenylation in vitro and in vivo (13, 44, 45, 49). Intensive biochemical studies have revealed a variety of *trans*-acting factors involved in the cleavage and polyadenylation reactions, including specificity factors, cleavage factors, and a poly(A) polymerase (7, 17, 20, 40, 42, 50).

Signals for mRNA 3′-end formation in plants and yeast cells appear to be different from those in mammals. Some mammalian polyadenylation sites have been shown to be nonfunctional in plants (21). In both plant and yeast cells, the essential sequences are situated distinctly further upstream of the cleavage site than in mammals. These findings indicate that these organisms might share mechanistic similarities in 3′-end formation that are possibly related to rho-dependent termination in bacteria (32).

Only a few plant polyadenylation sites have been characterized. Deletion analyses of the signals of the pea ribulose-1,5-biphosphate carboxylase small-subunit gene (*rbcS*) and of the cauliflower mosaic virus (CaMV) have revealed that sequences 30 to 150 nucleotides upstream of the cleavage site are essential for 3′-end formation (28, 38). The universal mammalian signal sequence AATAAA is present in only 40% of plant genes (25), but a deletion of this sequence in the case of CaMV abolishes 3′-end formation. However, the observation that point mutations that inactivate mammalian polyadenylation signals (44) result in only partial inactivation in plants suggests that a putative factor recognizing this sequence would not possess such sequence specificity as in mammals (28, 38).

In *Saccharomyces cerevisiae*, it has been shown that many mRNA 3′ ends are formed by endonucleolytic cleavage and subsequent polyadenylation analogous to that in higher eukaryotes (4, 5), although most genes lack the canonical AATAAA motif. In the *ADH2* gene, this motif does not act as a signal sequence (22). No unique sequence element has been identified near yeast poly(A) addition sites. A comparison of various polyadenylation sites suggests that in yeast cells, at least two different classes of polyadenylation sites exist (23). Whereas one class contains no significant sequence similarities, a second class contains the tripartite sequence TAG...TA(T)GT..TTT, which was first suggested to be a yeast 3′-end-formation signal by Zaret and Sherman (47) a decade ago. Mutagenesis of this sequence located 80 to 120 bp upstream of the *ARO4* poly(A) addition site dramatically reduced the activity of this site (23). Other recent reports have supported the involvement of this or a similar sequence in 3′-end formation. Single point mutations that inactivated the *ADH2* polyadenylation site have all been localized 70 to 90 bp upstream of the cleavage site within a sequence with high homology to the tripartite consensus sequence (22). Several intragenic revertants of a *cyc1-512* mutant defective in 3′-end formation were isolated and were shown to restore correct 3′-end formation because they had acquired either the bipartite motif TAG...TATGTA, a se-

* Corresponding author.
† Present address: Agriculture Canada Research Station, Vancouver, British Columbia V6T 1X2 Canada.

quence with high similarity to the previously proposed tripartite sequence, or an alternating A+T stretch (36, 48). Both sequences were proposed to act as 3'-end-formation signals in yeast cells. The involvement of an A+T alternating stretch was supported by the analysis of the signals for 3'-end formation in the *GAL7* gene. A 26-bp region encompassing the sequence $(AT)_9$ was shown to direct cleavage and polyadenylation 4 to 5 bp downstream of this gene (1).

In this report, we aimed to investigate putative similarities of 3'-end formation in plants and yeast cells. We demonstrate that *S. cerevisiae* specifically recognizes the polyadenylation signal of CaMV, a plant pararetrovirus (9, 14). The mRNA 3' endpoints, where the poly(A) tail is added, are at the same position as in plants, and the functioning of this site in yeast cells is orientation dependent. We show that multiple *cis*-acting sequence elements, including two upstream sequences, the AATAAA motif, and the sequences at or near the poly(A) addition site, are required for proper CaMV 3'-end formation in yeast cells. As the key element for function of the CaMV polyadenylation site in yeast cells, we have identified the sequence TAGTATGTA, a motif resembling previously proposed consensus sequences (36, 47). Point mutations in this sequence reduced the activity of this site to 10 to 20% of the original activity.

MATERIALS AND METHODS

Plasmid constructions. The plasmid pME729 that was used for inserting and analyzing the CaMV polyadenylation site and its various mutants was constructed as follows. The 0.45-kb *Hin*dIII-*Bam*HI fragment of pAAH5 (2) containing the 3'-end region of the *ADH1* gene was inserted into the polylinker of the *Escherichia coli-S. cerevisiae* shuttle vector YEp352 (19), a 2μm-derived plasmid containing the *URA3* gene for selection in *S. cerevisiae*, to create plasmid YEp352-ADH1. A 0.73-kb *Bss*HII-*Xho*I DNA fragment from plasmid pYactI (29) that contained the 5' region of the actin gene was blunt ended at the *Xho*I site and ligated into the *Nar*I-*Hin*dIII-digested vector YEp352-ADH1, which was filled in at the *Hin*dIII site, to create plasmid pME729. This plasmid therefore contains a fusion of the 5' part of the actin gene (promoter, 5' untranslated region, exon 1, and part of the intron) to the *ADH1* 3'-end region. The single *Xho*I site in between was used for inserting additional fragments. From the actin part, 0.22 kb are transcribed, and from the *ADH1* part, 0.22 kb are transcribed. The CaMV 3' end was received as a 0.250-kb *Sma*I-*Hin*dIII fragment from plasmid pDH51 (31) and was placed blunt ended into the filled-in *Xho*I site of pME729 to create pME729-CaP and pME729-CaN, with positive or negative orientation of the CaMV fragment to the direction of transcription, respectively. Large deletions were made by incubation of plasmid pDH51, which was linearized with *Pst*I (for 5' deletions) or *Hin*dIII (for 3' deletions), with nuclease *Bal* 31 for various periods. Linkers were ligated, and the shortened CaMV sequences were cloned at the site of the full-length CaMV fragment into pME729-CaP. Small deletions and point mutations were created with oligonucleotides containing the desired mutations by the polymerase chain reaction according to the method of Giebel and Spritz (15). These products were inserted at the place of the normal CaMV sequence into pME729-CaP. All deletions and mutations were confirmed by DNA sequencing.

RNA analysis. Total RNA from yeast cells was isolated according to the method of Zitomer and Hall (51) by using glass beads to disrupt the yeast cells. For Northern (RNA) hybridization, 10 μg of total RNA was separated on denaturing formaldehyde gels, transferred to nylon membranes, cross-linked by UV light, and hybridized with DNA fragments labeled according to the oligolabeling method described by Feinberg and Vogelstein (12). Band intensities on autoradiograms were quantitated with a densitometer (Macintosh). Nuclease S1 protection experiments were performed as previously described (23), and RNase protection analysis was performed according to the method of Goodall and Filipowicz (18).

Yeast strains and yeast methods. The *S. cerevisiae* strain used for this study was RH1376 (*MAT***a** *Δura3*), a derivative of the laboratory strains X2180-1A (*MAT***a** *gal2 SUC2 mal CUP1*) and X2180-1B (*MATα gal2 SUC2 mal CUP1*). Yeast cells were grown on YEPD (yeast extract-peptone-dextrose) complete medium or on MV minimal medium (27). *S. cerevisiae* transformation (24) and total DNA isolation (3) were previously described. Plasmid-carrying yeast strains were identified by Southern hybridization (41).

Reagents. Restriction enzymes, nucleases *Bal* 31 and S1, DNA and RNA polymerases, and other enzymes were purchased from Boehringer (Mannheim, Germany), Pharmacia (Uppsala, Sweden), or New England BioLabs (Schwalbach, Germany). Oligonucleotides were synthesized by Microsynth (Windisch, Switzerland).

RESULTS

Plant CaMV polyadenylation site is functional in yeast cells. We have constructed a simple test system for the analysis of 3'-end-formation signals in *S. cerevisiae*. This system is based on a 2μm-derived plasmid, pME729, which contains the 5' region of the actin gene, including its efficient promoter, exon 1, 80 bp of the intron, and, adjacent to it, the polyadenylation site of the *ADH1* gene. With this system, we have tested the polyadenylation site of a plant pararetrovirus, CaMV, which was inserted as a 250-bp fragment in both orientations into the single *Xho*I restriction site between the actin and *ADH1* sequences (Fig. 1A). The sequence of the CaMV 3'-end region is shown in Fig. 1B. The 250-bp CaMV fragment contains the region 180 bp upstream and 15 bp downstream of the poly(A) addition site in plants as well as polylinker sequences on both sides. Readthrough transcription from the actin promoter through the CaMV insert resulted in a transcript of 750 bp which was polyadenylated at the efficient *ADH1* polyadenylation site (23). 3'-End formation within the CaMV sequence produced a truncated RNA of about 500 bp (Fig. 1A). By using this test system, it was possible to quantitate the efficiency of 3'-end formation by determining the ratio of truncated versus readthrough transcripts.

Total RNA of the yeast strains containing these constructs was analyzed by Northern hybridization employing a DNA fragment from the actin region as probe. The results show that the CaMV polyadenylation site is recognized in an orientation-dependent manner (Fig. 2). When the CaMV sequence was inserted in a forward orientation into the test system, approximately 60% of all transcripts initiating from the actin promoter were truncated, whereas with reverse insertion, the amount of truncated transcript was negligible. Similar results were obtained with poly(A)$^+$ RNA enriched on oligo(dT) columns, indicating that all transcripts were properly polyadenylated (data not shown).

mRNA 3' ends produced by the CaMV polyadenylation site are identical in plants and yeast cells. The precise mRNA 3' ends of the truncated transcripts were determined by RNase

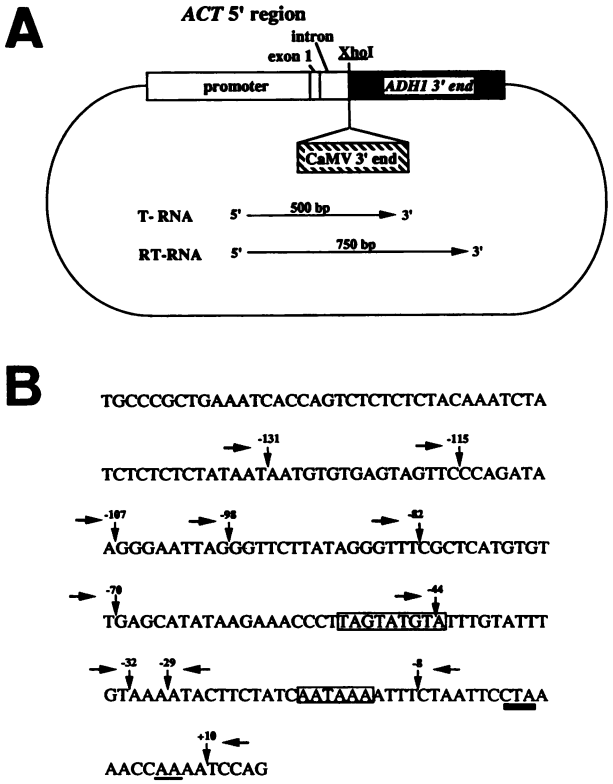

FIG. 1. (A) Test system for the CaMV polyadenylation signal in yeast cells. The CaMV fragment was inserted in both orientations into plasmid pME729 between the 5′ region of the actin gene (promoter, 5′ untranslated region, exon 1, and part of the intron) and the 3′ region of the *ADH1* gene, including its polyadenylation signal. Some essential sequences for splicing in the actin intron were missing. The readthrough RNA (RT-RNA) from the actin promoter through the CaMV sequence to the *ADH1* polyadenylation site and the truncated RNA (T-RNA) with 3′ ends in the CaMV sequences are illustrated with arrows. (B) Partial CaMV DNA sequence (14) encoding the polyadenylation signal. Underlined with a thick line are the CaMV major mRNA 3′ endpoints mapped in plants and yeast cells, and underlined with a thin line is an additional minor 3′ end found in yeast cells. The sequence motifs AATAAA at positions −12 to −17 with respect to the major poly(A) addition site and TAGTATGTA at positions −43 to −51 are boxed and are discussed in the text. Vertical arrows indicate the endpoints of various 5′ and 3′ deletions which were tested. The numbers represent the distances from deletion endpoints to the major poly(A) addition site. Horizontal arrows show the directions of the deletions that were performed from restriction sites upstream or downstream of the poly(A) addition site.

protection assays and nuclease S1 mappings and were compared with the mRNA 3′ ends mapped in transiently transfected plant protoplasts of *Nicotiana plumbaginifolia* (9, 38).

By using the RNase protection assay and a CaMV-specific, radiolabeled antisense probe, it was possible to directly compare mRNA isolated from plants and yeast cells. The results show that the major signals were located at the same positions in both organisms. Some minor signals are at different locations and might be due to nuclease artifacts (Fig. 3A).

In addition, nuclease S1 mapping experiments were performed with RNAs from yeast strains containing the forward and reverse inserted CaMV polyadenylation sites (Fig. 3B). The results obtained with this method confirmed the RNase

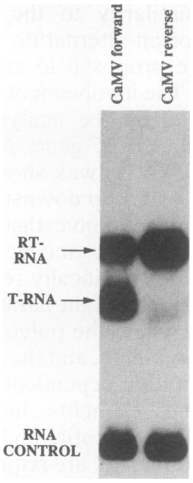

FIG. 2. Effect of the CaMV polyadenylation signal on mRNA 3′-end formation in yeast cells. Total RNA (10 μg) from yeast strains carrying the test plasmid with the CaMV polyadenylation signal inserted in forward or reverse orientation was separated on a formaldehyde-containing 2% agarose gel, transferred to nylon filters, and hybridized with a radiolabeled 450-bp *Mlu*I-*Xho*I *ACT* DNA fragment spanning the 5′ untranslated region, exon 1, and intron sequences (29). The readthrough transcript of about 750 bp (RT-RNA) and the truncated transcript of approximately 500 bp (T-RNA) are indicated. As a standard for the amount of RNA loaded, the same filter was washed and hybridized with a 1.1-kb *Hin*dIII-*Hin*dIII *URA3* DNA fragment (RNA control).

protection results. In the case of forward insertion, the major signal of the 3′ end of the truncated RNA was exactly at the same position as in plants. An additional minor signal was located 8 bp downstream of the major site (indicated in Fig. 1B). Additional protected bands that correspond to the readthrough RNA were found at the *ADH1* poly(A) addition site. The ratio of the truncated versus the readthrough RNA is significantly lower than in the Northern and RNase protection experiments. This observation is probably due to the higher affinity of the hybridization probe for the readthrough RNA.

At least four sequence elements are required for efficient function of the CaMV polyadenylation site in yeast cells. The specific recognition of the CaMV polyadenylation site in yeast cells and the identities of the mRNA 3′ endpoints prompted us to ask whether yeast cells require the same *cis*-acting elements as plants. In addition to the AATAAA motif, upstream elements have previously been shown to be important for proper function of the CaMV polyadenylation site in plants (28, 38).

In order to locate the essential sequences in yeast cells, a deletion set was created by using nuclease *Bal* 31. In eight *Bal* 31 deletions, the sequences from the *Pst*I restriction site at the 5′ end of the polyadenylation signal toward the poly(A) addition site were progressively removed (5′ deletions). The endpoints of these deletions are at positions −131, −115, −107, −98, −82, −70, −44, and −32 with respect to the poly(A) addition site (indicated in Fig. 1B). In three additional deletions, the sequences from the *Hin*dIII restriction site situated at the 3′ end of the CaMV sequence toward the polyadenylation signal were deleted (3′ deletions). The endpoints of these deletions are at positions +10, −8, and −29 relative to the poly(A) addition site. The various mutant CaMV polyadenylation sites were cloned

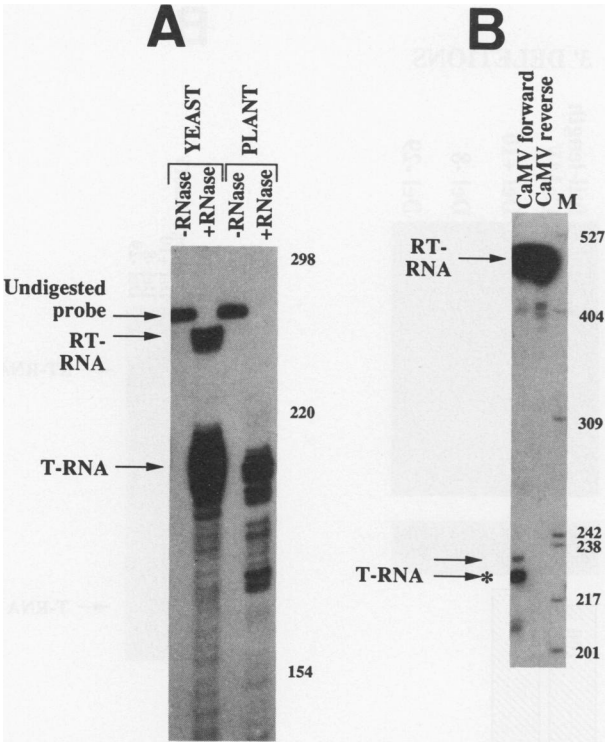

FIG. 3. (A) RNase protection assays. Total RNA from the yeast strain carrying the test plasmid with the forward inserted CaMV polyadenylation site and total RNA from transiently transfected plant protoplasts (*N. plumbaginifolia*) were hybridized with the same CaMV antisense RNA probe (39). The hybrids were digested with RNase A and T1. The undigested RNA probe, the readthrough RNA (RT-RNA), the truncated RNA (T-RNA), and the sizes of *Hin*fI-digested pBR322 size markers are illustrated. (B) Nuclease S1 protection. Total RNA (20 μg) of the yeast strains containing the test plasmid with the forward or reverse inserted CaMV fragment were hybridized with a complementary 3′-end-labeled DNA fragment ranging from the *Ava*II site situated in the actin intron 15 bp upstream of the CaMV insert to the *Sma*I site in the polylinker downstream of the *ADH1* fragment. The asterisk represents the mRNA 3′ end that corresponds to the mRNA 3′ end mapped in plants. The positions of the 3′ ends of the truncated RNA (T-RNA) in the CaMV fragment are shown in Fig. 1B. The readthrough transcripts (RT-RNA) end at the *ADH1* polyadenylation site. The sizes of the signals were determined by using the *Hpa*II-digested plasmid pBR322 (lane M) and G+A sequencing ladders (not shown).

into the test plasmid, and 3′-end-formation efficiency was determined by Northern analysis (Fig. 4A). The results demonstrate that several elements are important for 3′-end formation in yeast cells. Deletions of sequences more than 131 and 115 bp upstream of the mapped mRNA 3′ endpoints reduced the activity to about 60 and 30%, respectively, compared with that of the undeleted fragment. Several deletions to positions −107, −98, −82, and −70 did not further decrease the activity significantly. Additional deletions, however, removing the sequences to positions −44 and −32 upstream of the poly(A) addition site, abolished 3′-end formation. When the RNA was analyzed in nuclease S1 protection assays, we found that the deletions did not alter the accuracy of the mRNA 3′ endpoints (data not shown).

Removal of sequences more than 10 bp downstream of the poly(A) addition site (Del +10 in Fig. 4A) did not reduce the efficiency of 3′-end formation, but the deletion of the normal poly(A) addition site (Del −8) caused a loss of activity to 30%, and an additional deletion of the region encompassing the sequence motif AATAAA (Del −29) reduced the efficiency to below 10%. Nuclease S1 protection assays showed that deletion of the normal poly(A) addition site resulted in the production of several heterogeneous 3′ ends downstream of the normal site, and the additional deletion of the region encompassing the sequence AATAAA reinforced this effect (Fig. 4B).

In summary, this deletion analysis suggests that for proper function of the CaMV polyadenylation signal, at least four sequence elements are required, namely, a distal (>115 bp) and a proximal (−44 to −70) upstream element, the region containing the AATAAA sequence, and the region encompassing the poly(A) addition site.

Deletion of the AATAAA motif or mutagenesis to AAGAAA reduces activity of the CaMV polyadenylation site in yeast cells to 50%. The sequence AATAAA is located 12 to 17 bp upstream of the CaMV poly(A) addition site. Deletion of this sequence inactivated the polyadenylation site in plants, while point mutations reduced its function to 50% (28, 38).

Our deletion analysis showed that the region containing the AATAAA motif is required in yeast cells. We therefore analyzed the effects of mutant CaMV polyadenylation sites containing either a complete deletion of the AATAAA motif or a point mutation to AAGAAA. Northern hybridizations (Fig. 5) demonstrated that both mutations had distinct effects on the efficiency of 3′-end formation in yeast cells. The amount of truncated RNA was reduced, and more readthrough transcripts were produced. Quantification of the band intensities revealed that both mutations decreased the efficiency of the site to about 50% compared with that of the normal CaMV polyadenylation site. Mapping of the mRNA 3′ ends indicated that although the formation of truncated RNA was lowered, the locations of the mRNA 3′ endpoints were not altered (data not shown).

The sequence TAGTATGTA is the key element for proper function of the CaMV polyadenylation site in yeast cells. Deletion analysis had indicated that the deletion of the element located between −44 and −70 completely abolished 3′-end formation (Fig. 4A). This region contains at positions −43 to −51 the DNA sequence TAGTATGTA, which is similar to a bipartite sequence (TAG...TATGTA) that was previously proposed as a 3′-end-formation signal sequence (36, 47). Interestingly, the major upstream element in plants, localized between positions −32 and −44, is situated immediately adjacent to the TAGTATGTA motif, which has, however, been shown to be dispensable for CaMV 3′-end formation in plants (38).

In order to characterize the sequence elements between positions −44 and −70, we constructed two small deletion mutants in this region. One of them lacked the sequences from −53 to −70 (Δ−53−70), therefore leaving the TAG TATGTA motif intact. In the second deletion, from positions −46 to −56 (Δ−46−56), this motif was destroyed. Analysis of RNAs from the deletion mutants revealed that mutant Δ−53−70 even had an increased activity of the polyadenylation site, whereas mutant Δ−46−56 had a dramatically decreased activity, i.e., below 5% relative to that of the wild-type CaMV sequence (Fig. 6).

We introduced point mutations into the sequence TAG TATGTA, namely, 1-, 2-, and 3-bp alterations that changed the sequence to TAGCATGTA, TCGCATGTA, and TCG CATTTA, respectively. The effects of the mutations were analyzed by Northern hybridization (Fig. 6). The single point

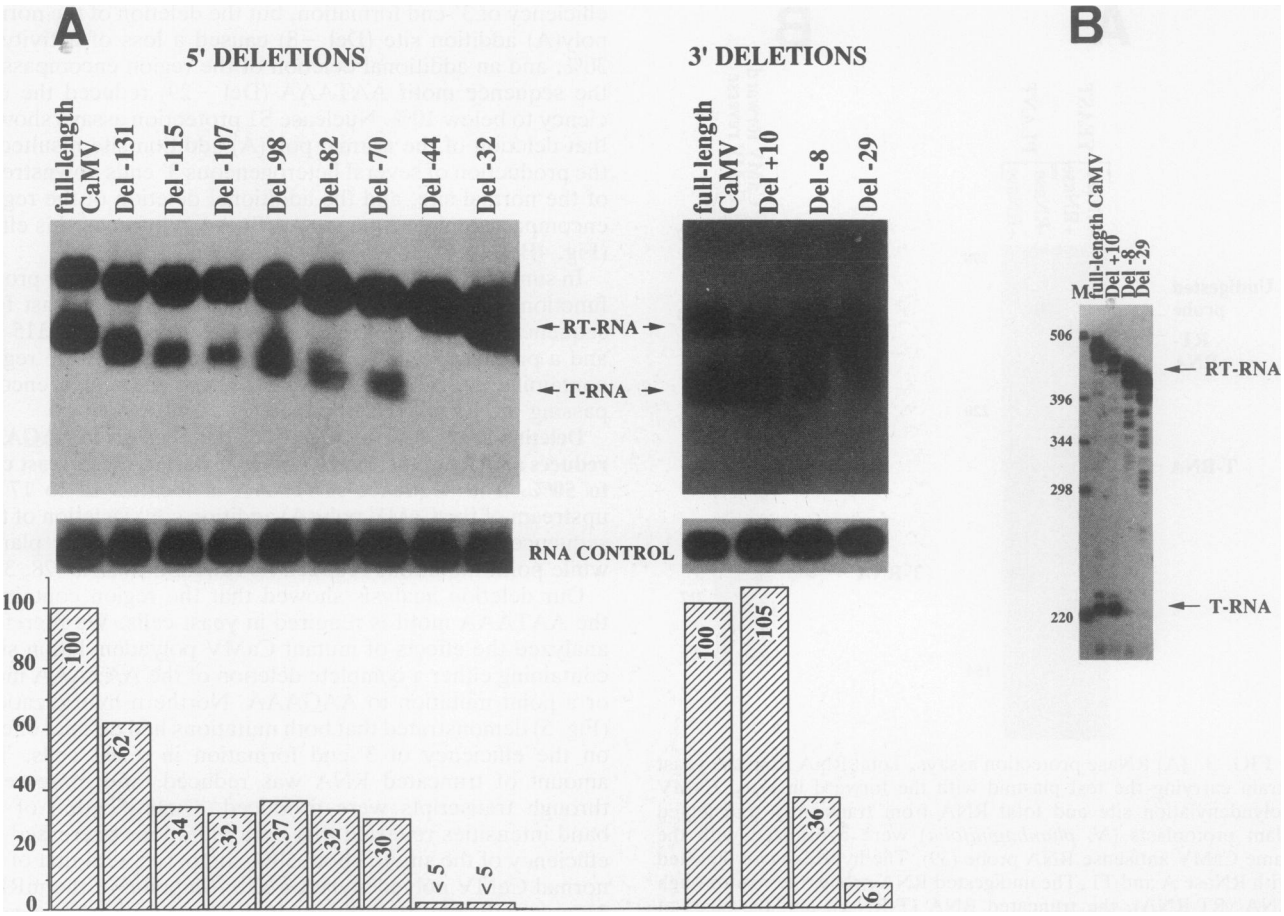

FIG. 4. Deletion analysis of the CaMV polyadenylation signal in yeast cells. (A) Northern hybridization with total RNA from yeast strains carrying plasmids with partially deleted CaMV sequences. Northern hybridization experiments were performed as described in the legend to Fig. 2. 5′ Deletions from the *Pst*I restriction site upstream of the polyadenylation signal are shown on the left, and 3′ deletions from the *Hin*dIII restriction site downstream of the polyadenylation signal are shown on the right. Numbers correspond to the deletion endpoints and represent the distance from the major poly(A) addition site (Fig. 1B). Quantitation of the Northern experiments is depicted below. The intensities of the truncated-RNA (T-RNA) and readthrough RNA (RT-RNA) bands were evaluated with a densitometer. Each quantitation result corresponds to an average value from at least six independent Northern blotting experiments. The standard deviation did not exceed 20%. The 3′-end-formation efficiencies for the mutants in relation to that for the wild-type CaMV sequence are indicated. (B) Nuclease S1 protection. Determination of the mRNA 3′ ends of the 3′ deletions of the CaMV polyadenylation signal is shown. The experiments were performed as described in the legend to Fig. 3B by using the *Hin*fI-digested plasmid pBR322 (lane M) as a size standard.

mutation at position −48 reduced 3′-end formation in the CaMV fragment to 10 to 20% of the original amount; the multiple mutations caused a reduction to below 10% of that of the CaMV wild-type sequence. In all cases, increased readthrough transcription was observed, suggesting that the mutations more likely affected 3′-end formation rather than mRNA stability.

Since the sequence TAGTATGTA was suggested to act in yeast cells as a bipartite 3′-end-formation signal, TAG... TATGTA (36), we tested the influence of a mutation that introduces spacing into this sequence. The results of this mutation demonstrate that introduction of 4 bp into the signal sequence, creating the sequence TAGatatTATGTA, increased the efficiency of the polyadenylation site to 140% compared with that of the normal CaMV sequence (Fig. 6). The amount of truncated transcripts is enhanced and readthrough transcription is lowered, so that now 85% of the transcripts initiating from the actin promoter end at the CaMV 3′ end, in comparison with 60% in the wild-type sequence. This mutation affected only the efficiency of the signal but not the accuracy of the mRNA 3′ ends as determined by nuclease S1 protection (data not shown). The results of the mutational analysis of the CaMV 3′-end-formation signal are summarized in Fig. 7, where the sequence elements we identified as important in yeast cells are illustrated and compared with the essential elements identified in plants (according to reference 38).

DISCUSSION

We have shown in this paper that the polyadenylation signal of CaMV is specifically recognized in *S. cerevisiae*. The CaMV mRNA 3′ endpoints and the positions where the poly(A) tail is added are identical in yeast cells and plants. The specific function is characterized by the orientation-dependent function: with forward orientation, approximately 60% of the transcripts from an efficient yeast promoter were polyadenylated at the CaMV poly(A) site, whereas reverse insertion of the polyadenylation site into a test plasmid did not significantly cause 3′-end formation. The

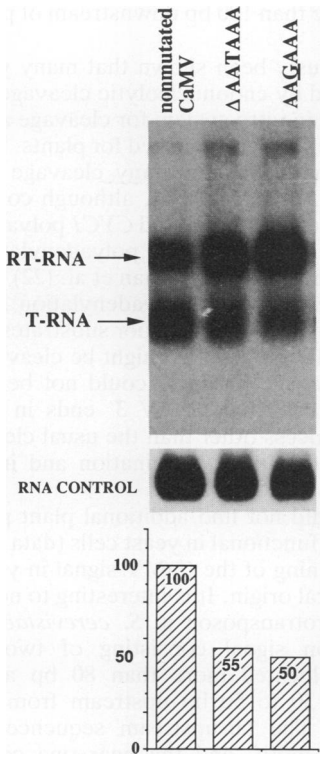

FIG. 5. Effects of mutations of the CaMV AATAAA motif in yeast cells. Either the sequence AATAAA at positions −12 to −17 (compare with Fig. 1B) was deleted (ΔAATAAA), or a single point mutation was created (AAGAAA). Total RNAs of yeast strains carrying the mutated plasmids were analyzed by Northern blotting as described in the legend to Fig. 2. The ratios of the truncated (T-RNA) versus the readthrough (RT-RNA) RNAs were quantitated as explained in the legend to Fig. 4 and are shown in the diagram at the bottom. The relative efficiencies of 3′-end formation of the mutated CaMV sequences in comparison with those of the nonmutated sequence are indicated.

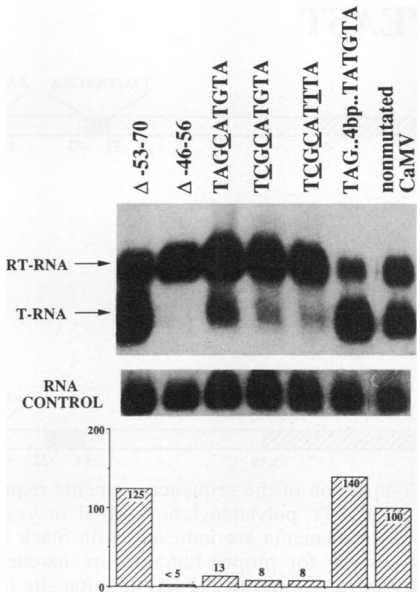

FIG. 6. Effects of deletions and point mutations in the region −44 to −70 upstream of the CaMV poly(A) addition site. The CaMV sequence was mutagenized by using oligonucleotide-directed mutagenesis procedures: the sequences from −53 to −70 (Δ−53−70) and from −46 to −56 (Δ−46−56) with respect to the poly(A) addition site were deleted, and the consensus sequence, TAGTATGTA, located between nucleotides −43 and −51, was altered by 1-bp (TAGCATGTA), 2-bp (TCGCATGTA), and 3-bp (TCGCATTTA) mutations and by a 4-bp insertion (TAGatatTAGTATGTA) (lowercase letters indicate inserted nucleotides). Altered nucleotides are underlined. Northern hybridization experiments were performed as described in the legend to Fig. 2 and quantitated as described in the legend to Fig. 4. The quantitation of the ratios of the truncated (T-RNA) versus the readthrough (RT-RNA) transcripts is depicted in the diagram at the bottom. The values correspond to the relative 3′-end-formation efficiency compared with that of the nonmutated sequence.

efficiency of this site is comparable with that of three yeast polyadenylation sites (*ARO4*, *TRP1*, and *TRP4*) that were previously shown to be less efficient than other polyadenylation sites (23).

Defining the *cis* elements required for recognition of the CaMV polyadenylation signal in yeast cells allowed us to compare the sequence requirements in yeast cells and plants. In plants, point mutations of the AATAAA sequence cause a partial reduction, but complete deletion of this motif abolishes 3′-end formation (28, 38). Deletion of far-upstream sequences, more than 100 bp upstream of the poly(A) addition site, reduce the activity to 60%, and a more-important near-upstream element is localized between positions −32 and −44 (38).

We have shown that in yeast cells, at least two upstream elements are required for proper function of the CaMV polyadenylation signal. Mutational analysis of the CaMV polyadenylation signal in yeast cells revealed that deletion of far-upstream sequences located more than 115 bp upstream from the mRNA 3′ endpoint reduced the activity of the CaMV signal to about 30% of the original activity. A proximal upstream element that was indispensable was localized between positions −43 and −51, which is immediately adjacent to the essential sequences in plants. Interestingly, a deletion of 18 bp between the distal and proximal upstream elements enhanced the CaMV signal, possibly by moving the distal upstream sequences closer to the poly(A) addition site.

The proximal upstream element is the major element for recognition of the CaMV 3′-end-formation signal in yeast cells. A 10-bp deletion in this region completely abolishes 3′-end formation. This deletion removes the first 6 bp of the sequence motif TAGTATGTA, which was previously proposed to act as a bipartite signal sequence (TAG...TATGTA) in yeast cells (36). We showed by mutational analysis that this motif is indeed the key element for function of the CaMV signal in yeast cells, because one single point mutation reduced the activity to below 15%, whereas an insertional mutation that separated the CaMV motif to a bipartite sequence increased 3′-end formation at the CaMV polyadenylation site. Several recent reports provided evidence either for or against the importance of the TAG...TATGTA motif or similar sequences in yeast mRNA 3′-end formation. It was shown that in a *cyc1-512* mutant defective in 3′-end formation, mRNA deficiency was restored in several intragenic revertants that formed this sequence by a few base changes (36). Mutations in this sequence caused a drastic loss of activity at the *ARO4* polyadenylation site (23), and in the *ADH2* gene, single point mutations that inactivated the polyadenylation site were identified as situated in similar sequences (22). On the other hand, a mutational analysis of

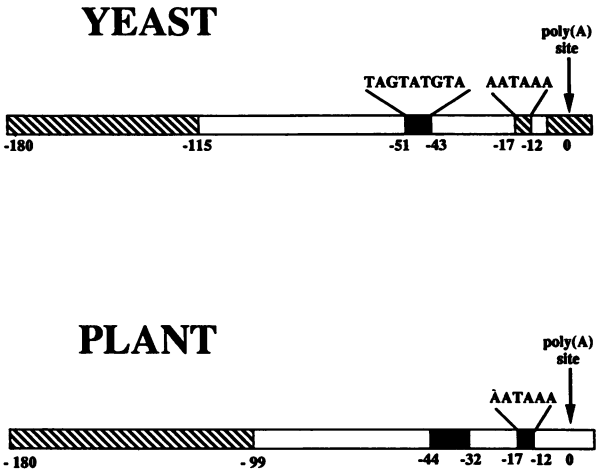

FIG. 7. Comparison of the sequence elements required for recognition of the CaMV polyadenylation signal in yeast cells and plants. The major elements are indicated with black boxes; other sequences necessary for proper function are hatched. Numbers represent the distance from the poly(A) addition site (marked with an arrow).

the *CYC1* polyadenylation site revealed that mutagenesis of the bipartite motif did not affect 3'-end formation (30). Our results clearly show that the sequence TAG...TATGTA is a signal for mRNA 3'-end formation in yeast cells and that this signal acts more efficiently when present as a bipartite sequence. They also show that this element acts efficiently only in concert with other elements. Similar results have been obtained with the *CYC1* polyadenylation site (36). Obviously, the bipartite sequence is only one of two or more signal sequences in yeast cells (36) which might be required for mRNA 3'-end formation in only one class of yeast genes (23).

Sequences more than 10 bp downstream of the CaMV poly(A) addition site are apparently not required, but elements close to the mRNA 3' endpoint are important for efficient function as well as for the accuracy of the mRNA 3' ends in yeast cells. These observations are in accordance with other reports that suggested the involvement of elements acting at or near the site where the poly(A) tail is added (1, 36, 46).

Mutagenesis of the AATAAA motif upstream of the mRNA 3' endpoint has distinct effects in yeast cells. This result was not expected, since this sequence is absent in most yeast genes and was shown to be not essential for the *ADH2* polyadenylation signal (22). For the CaMV signal, deletion of the AATAAA sequence as well as a single point mutation altering the T to a G caused a 50% loss of activity of the polyadenylation site. A possible explanation for these observations is that, in contrast to higher eukaryotic cells, the AATAAA does not act as a signal sequence for a protein factor but rather is required as a part of an A+T-rich region near the mRNA 3' endpoint that might act as an RNA polymerase II pause site. Both deletion of six A+T residues and introduction of a G residue might impair this region and cause it to function as a pause site. DNA regions with an A+T content of up to 80% are common in the 3' regions of many yeast genes, and such unusual DNA regions have been suggested as slowing down or stopping the elongation of RNA polymerase II (11). In accordance with such a model is the observation that transcription termination in yeast cells occurs not more than 100 bp downstream of poly(A) addition sites (37).

It has previously been shown that many yeast mRNA 3' ends are formed by endonucleolytic cleavage of a precursor RNA (1, 5). An in vitro system for cleavage and polyadenylation has not yet been established for plants. Using yeast cell extract, we could not detect any cleavage of an in vitro-synthesized CaMV pre-mRNA, although control yeast pre-mRNAs spanning the *ADH1* and *CYC1* polyadenylation sites were properly processed and polyadenylated in our cell extracts (data not shown). Hyman et al. (22) have found that pre-mRNAs of two yeast polyadenylation sites that functioned efficiently in vivo were poor substrates for cleavage in vitro. The CaMV pre-mRNA might be cleaved very slowly, so that the cleavage products could not be identified. Another possibility is that CaMV 3' ends in yeast cells are formed by a process other than the usual cleavage reaction, perhaps by transcription termination and immediate polyadenylation.

Since we could not find additional plant polyadenylation signals that are functional in yeast cells (data not shown), the specific functioning of the CaMV signal in yeast cells might be due to its viral origin. It is interesting to note that the Ty1 element, a retrotransposon in *S. cerevisiae*, has a similar 3'-end-formation signal consisting of two upstream sequences, one located more than 80 bp and the second approximately 30 to 50 bp upstream from the mRNA 3' endpoint (46). The far-upstream sequence has a perfect TAGTATGTA motif, and the near one contains a slight deviation from it. The sequence similarities of the CaMV and the Ty polyadenylation sites might explain the recognition of the plant viral signal in yeast cells. CaMV and Ty element are both genetic elements that replicate via reverse transcription of an RNA intermediate. The CaMV and Ty transcripts are analogous, since both contain polyadenylation signals at their 5' and 3' ends, similar to the situation in retroviruses, where the polyadenylation sites are duplicated in the 5' and the 3' long terminal repeats (for a review, see reference 33). The polyadenylation signals of retroelements in yeast cells and plants might have a common evolutionary origin. In this context, it will be interesting to analyze the characteristics of other plant viral signals in yeast cells as soon as data on such polyadenylation signals are available.

Although the functioning of plant polyadenylation sites in yeast cells seems not to be a general process, our results presented here confirm previous speculations that the polyadenylation signals in yeast cells and plants have some common characteristics. In both organisms, the essential *cis*-acting signal sequences are situated further upstream of the poly(A) addition site than in other eukaryotes. These observations suggest that yeast cells and plants both possess upstream recognition factors that require sequences far away from the cleavage site. The different sequence requirement in the CaMV polyadenylation site in yeast cells and plants suggests that these factors recognize different signal sequences in yeast cells and in plants. Analogous factors might also recognize the sequence elements upstream of the canonical AATAAA motif that are involved in 3'-end formation at the polyadenylation sites of various animal viruses, for example, in the adenovirus late transcript (10), the simian virus 40 late transcript (6), hepatitis B virus (35), and human immunodeficiency virus (43).

In future, it will be interesting to further evaluate functional similarities in mRNA 3'-end formation between yeast cells and plants, for example, by introducing yeast polyadenylation sites into plants, and to elucidate the similarities and

differences in the mechanisms of mRNA 3'-end formation among higher and lower eukaryotic organisms, viruses, and bacteria.

ACKNOWLEDGMENTS

We appreciate Ralf Hütter's generous support, and we are grateful to David Jones, Thomas Hohn, and Helen Rothnie for critical comments on the manuscript. We thank Christoph Kündig for synthesizing some oligonucleotides and Hans-Ulrich Mösch, Markus Künzler, and Roney Graf for helpful advice and suggestions.

This work was supported by the Swiss National Science Foundation, grant 31.29926.90, and the Swiss Federal Institute of Technology.

REFERENCES

1. **Abe, A., Y. Hiraoka, and T. Fukasawa.** 1990. Signal sequence for generation of mRNA 3' end in the *Saccharomyces cerevisiae GAL7* gene. EMBO J. **9**:3691–3697.
2. **Ammerer, G.** 1983. Expression of genes in yeast using the *ADC1* promoter. Methods Enzymol. **101**:192–201.
3. **Braus, G., R. Furter, F. Prantl, P. Niederberger, and R. Hütter.** 1985. Arrangements of genes *TRP1* and *TRP3* of *Saccharomyces cerevisiae* strains. Arch. Microbiol. **142**:383–388.
4. **Butler, J. S., and T. Platt.** 1988. RNA processing generates the mature 3' end of yeast *CYC1* mRNA *in vitro*. Science **242**:1270–1274.
5. **Butler, J. S., P. P. Sadhale, and T. Platt.** 1990. RNA processing in vitro produces mature 3' ends of a variety of *Saccharomyces cerevisiae* mRNAs. Mol. Cell. Biol. **10**:2599–2605.
6. **Carswell, S., and J. C. Alwine.** 1989. Efficiency of utilization of the simian virus late polyadenylation site: effect of upstream sequences. Mol. Cell. Biol. **9**:4248–4258.
7. **Christofori, G., and W. Keller.** 1988. 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor and a snRNP. Cell **54**:875–889.
8. **Conway, L., and M. Wickens.** 1985. A sequence downstream of AAUAAA is required for formation of simian virus 40 late mRNA 3' termini in frog oocytes. Proc. Natl. Acad. Sci. USA **82**:3949–3953.
9. **Covey, S. N., G. P. Lomonosoff, and R. Hull.** 1981. Characterization of cauliflower mosaic virus DNA sequences which encode major polyadenylated transcripts. Nucleic Acids Res. **9**:6735–6747.
10. **DeZazzo, J. D., and M. J. Imperiale.** 1989. Sequences upstream of AAUAAA influence poly(A) site selection in a complex transcription unit. Mol. Cell. Biol. **9**:4951–4961.
11. **Enriquez-Harris, P., N. Levitt, D. Briggs, and N. J. Proudfoot.** 1991. A pause site for RNA polymerase II is associated with termination of transcription. EMBO J. **10**:1833–1842.
12. **Feinberg, A. P., and B. Vogelstein.** 1984. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137**:266–267. (Addendum.)
13. **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.
14. **Franck, A., H. Guilley, G. Jonard, K. Richards, and L. Hirth.** 1980. Nucleotide sequence of cauliflower mosaic virus DNA. Cell **21**:285–294.
15. **Giebel, L. B., and R. A. Spritz.** 1990. Site-directed mutagenesis using a double-stranded DNA fragment as a PCR primer. Nucleic Acids Res. **18**:4947.
16. **Gil, A., and N. J. Proudfoot.** 1984. A sequence downstream of the AAUAAA is required for rabbit β-globin mRNA 3'-end formation. Nature (London) **312**:473–474.
17. **Gilmartin, G. M., M. A. McDevitt, and J. R. Nevins.** 1988. Multiple factors are required for specific RNA cleavage at a poly(A) addition site. Genes Dev. **2**:578–587.
18. **Goodall, G. J., and W. Filipowicz.** 1989. The AU-rich sequences present in the introns of plant nuclear pre-mRNA are required for splicing. Cell **58**:473–483.
19. **Hill, J. E., A. M. Myers, T. J. Koerner, and A. Tzagoloff.** 1986. Yeast/*E. coli* shuttle vectors with multiple unique restriction sites. Yeast **2**:163–167.
20. **Humphrey, T., G. Christofori, V. Lucijanic, and W. Keller.** 1987. Cleavage and polyadenylation of messenger RNA precursors in vitro occurs within large and specific 3' processing complexes. EMBO J. **6**:4159–4168.
21. **Hunt, A. G., N. M. Chu, J. T. Odell, F. Nagy, and N.-H. Chua.** 1987. Plant cells do not properly recognize animal gene polyadenylation signals. Plant Mol. Biol. **8**:23–35.
22. **Hyman, L. E., S. H. Seiler, J. Whoriskey, and C. L. Moore.** 1991. Point mutations upstream of the yeast *ADH2* poly(A) site significantly reduce the efficiency of 3'-end formation. Mol. Cell. Biol. **11**:2004–2012.
23. **Irniger, S., C. M. Egli, and G. H. Braus.** 1991. Different classes of polyadenylation sites in the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. **11**:3060–3069.
24. **Ito, H., Y. Fukuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153**:163–168.
25. **Joshi, C. P.** 1987. Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. Nucleic Acids Res. **15**:9627–9640.
26. **McLauchlan, J., D. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. Nucleic Acids Res. **13**:1347–1368.
27. **Miozzari, G., P. Niederberger, and R. Hütter.** 1978. Tryptophan biosynthesis in *Saccharomyces cerevisiae*: control of the flux through the pathway. J. Bacteriol. **134**:48–59.
28. **Mogen, B. D., M. H. MacDonald, R. Graybosh, and A. G. Hunt.** 1990. Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. Plant Cell **2**:1261–1272.
29. **Ng, R., and J. Abelson.** 1980. Isolation and sequence of the gene for actin in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **77**:3912–3916.
30. **Osborne, B. I., and L. Guarente.** 1989. Mutational analysis of a yeast transcriptional terminator. Proc. Natl. Acad. Sci. USA **86**:4097–4101.
31. **Pietrzak, M., R. D. Shillito, T. Hohn, and I. Potrykus.** 1986. Expression in plant of two bacterial antibiotic resistance genes after protoplast transformation with a new plant expression vector. Nucleic Acids Res. **14**:5857–5868.
32. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
33. **Proudfoot, N.** 1991. Poly(A) signals. Cell **64**:671–674.
34. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3' non-coding region sequences in eukaryotic messenger RNA. Nature (London) **263**:211–214.
35. **Russnak, R., and D. Ganem.** 1990. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. Genes Dev. **4**:764–776.
36. **Russo, P., W.-Z. Li, D. M. Hampsey, K. S. Zaret, and F. Sherman.** 1991. Distinct *cis*-acting signals enhance 3' endpoint formation of *CYC1* mRNA in the yeast *Saccharomyces cerevisiae*. EMBO J. **10**:563–571.
37. **Russo, P., and F. Sherman.** 1989. Transcription terminates near the poly(A) site in the *CYC1* gene of the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **86**:8348–8352.
38. **Sanfaçon, H., P. Brodmann, and T. Hohn.** 1991. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. Nature (London) **346**:81–84.
39. **Sanfaçon, H., and T. Hohn.** 1990. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. Nature (London) **346**:81–84.
40. **Skolnik-David, H., C. L. Moore, and P. A. Sharp.** 1987. Electrophoretic separation of polyadenylation specific complexes. Genes Dev. **1**:672–682.
41. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
42. **Takagaki, Y., J. L. Manley, C. C. MacDonald, J. Wilusz, and T.**

Shenk. 1990. A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. Genes Dev. **4:**2112–2120.

43. **Valsamakis, A., S. Zeichner, S. Carswell, and J. C. Alwine.** 1991. The human immunodeficiency virus type 1 polyadenylation signal: a 3′ long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. Proc. Natl. Acad. Sci. USA **88:**2108–2112.
44. **Wickens, M.** 1990. How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. **15:**277–281.
45. **Wickens, M., and P. Stephenson.** 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3′ end formation. Science **226:**1045–1051.
46. **Yu, K., and R. T. Elder.** 1989. Some of the signals for 3′ end formation in transcription of the *Saccharomyces cerevisiae* Ty-D15 element are immediately downstream of the initiation site. Mol. Cell. Biol. **9:**2431–2444.
47. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28:**563–573.
48. **Zaret, K. S., and F. Sherman.** 1984. Mutationally altered 3′ ends of yeast *CYC1* mRNA affect transcript stability and translational efficiency. J. Mol. Biol. **176:**107–135.
49. **Zarkower, D., P. Stephenson, M. Sheets, and M. Wickens.** 1986. The AAUAAA sequence is required both for cleavage and for polyadenylation of simian virus 40 pre mRNA in vitro. Mol. Cell. Biol. **6:**2317–2323.
50. **Zarkower, D., and M. Wickens.** 1987. Specific pre-cleavage and post-cleavage complexes involved in the formation of SV40 late mRNA termini in vitro. EMBO J. **6:**177–186.
51. **Zitomer, R. S., and B. D. Hall.** 1976. Yeast cytochrome c messenger RNA. J. Biol. Chem. **251:**6320–6326.