MOLECULAR AND CELLULAR BIOLOGY, June 1996, p. 2772–2776 Vol. 16, No. 6
0270-7306/96/$04.00+0

Signals Sufficient for 3′-End Formation of Yeast mRNA

ZIJIAN GUO† AND FRED SHERMAN*

Department of Biochemistry, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642

Received 14 November 1995/Returned for modification 21 December 1995/Accepted 22 March 1996

The following three elements were previously shown to be required for 3′-end formation of mRNA in the yeast *Saccharomyces cerevisiae*: (i) the efficiency element TATATA or related sequences, which function by enhancing the efficiency of downstream positioning elements; (ii) the positioning element AATAAA or related sequences, which position the poly(A) site; and (iii) the actual poly(A) site, which is usually Py(A)$_n$. In this study, we synthesized a 39-bp poly(A) signal that contained the optimum sequences of these three elements. By inserting the synthetic 3′-end-forming signal into various positions of a *CYC1-lacZ* fusion gene, we showed that truncated transcripts of the expected sizes were generated. Furthermore, the poly(A) sites of the truncated transcripts were mapped to the expected poly(A) site within the synthetic signal. Our findings establish that the three elements are not only necessary but also sufficient for mRNA 3′-end formation in *S. cerevisiae*.

3′-end formation of higher eukaryotic mRNAs requires an invariant AAUAAA sequence approximately 20 nucleotides (nt) upstream from the poly(A) site (21). In contrast, 3′-end formation of mRNAs from the yeast *Saccharomyces cerevisiae* does not use highly conserved motifs (1, 3, 9, 23) but requires other signals that were first uncovered with the *cyc1-512* deletion (23). This *cyc1-512* deletion abolished normal 3′-end formation of the *CYC1* mRNA, resulting in low levels of *CYC1* mRNA with extended 3′ ends. Furthermore, the same 38-bp region of *CYC1* was found to be required for 3′-end processing of extended *CYC1* transcripts in vitro (2). Detailed mutational analyses of the 38-bp region (6, 16), as well as other studies (8, 10), indicated that the *CYC1* mRNA 3′-end-forming signal has functional redundancy and suggested that 3′-end formation of yeast mRNAs involves signals having the following three degenerate elements: (i) the efficiency element, which functions by enhancing the efficiency of positioning elements; (ii) the positioning element, which positions the poly(A) site; and (iii) the actual site of polyadenylation. Sequence requirements at the poly(A) site have been systematically investigated with both the *CYC1* and *ADH1* genes (8, 16), and the results suggested that the Py(A)$_n$ sequences, i.e., cytidine or thymidine followed by one or more adenosine residues, are preferential sites for polyadenylation. Studies from several groups (1, 6, 10, 17) have indicated that TATATA and related sequences function as efficiency elements and that TATATA is the most efficient of these. Recently, we characterized the positioning element and showed that its optimum sequence was AATAAA or AAAAAA (7). An efficiency element, TATATA, combined with a downstream AATAAA sequence, was found to direct 3′-end formation in vivo and 3′-end processing in vitro in a *CYC1* context (7).

The results from several other studies of various yeast genes are also consistent with the view that the three elements are necessary for 3′-end formation of mRNAs. In addition to the *CYC1* gene, sequence elements related to TATATA have been identified in several yeast genes and heterologous sequences, which appeared to determine the efficiency of 3′ end formation (1, 6, 9–11, 17, 22). A deletion within the *GCN4* 3′ untranslated region was found to result in a shift of poly(A) sites without decreasing the overall 3′-end formation efficiency (3). A sequence related to AATAAA within this *GCN4* deletion could function as a positioning element. The poly(A) sites of several yeast genes have been accurately mapped by sequencing of the corresponding cDNAs, and the results showed that they were all composed of a Py(A)$_n$ sequence (7).

Although it is clear that the three elements are involved in mRNA 3′-end formation, it has yet to be determined whether they are sufficient and whether other sequences are also required or influence 3′-end formation. In this study, we have constructed a synthetic oligonucleotide containing the optimum sequences of the three distinct elements, separated by random sequences. We demonstrated that this synthetic signal is sufficient to direct 3′-end formation in the transcription unit of a *CYC1-lacZ* fusion gene (5). As expected, the actual poly(A) sites were mapped to the sequence TTTCAAA within the synthetic signal. Because the sequence context in *CYC1-lacZ* is G+C rich, an A+T-rich context does not appear to be required for the function of the yeast 3′-end-forming signal.

MATERIALS AND METHODS

Yeast strain. *S. cerevisiae* B-7467 (*MAT***a** *cyc1-Δ1 CYC7*$^+$ *ura3-52*) was used as the recipient of the plasmids containing the *CYC1-lacZ* fusion gene and its derivative genes with the insertions of synthetic 3′-end-forming signal.

Oligonucleotides. Oligonucleotides used for construction of plasmids, Northern (RNA) blot analysis, and PCR mapping were synthesized on an Applied Biosystems 380A DNA synthesizer and are listed in Table 1.

Construction of plasmids. Plasmid pAA1598 (pLG-312) was kindly provided by L. Guarente (5). The synthetic poly(A) signal sequence was inserted into pAA1598 as duplex oligonucleotides (Fig. 1C). The following pairs of single-stranded oligonucleotides were annealed to each other in 1× SSC buffer (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) by being heated to 70°C for 2 min and cooled down to room temperature within 2 h: OL95-88 and OL95-89, OL95-91 and OL95-92, OL95-86 and OL95-90, and OL95-85 and OL95-87. Approximately 20 ng of the duplex oligonucleotides was ligated with 0.1 μg of pAA1598 digested with the pertinent restriction enzymes. Because the synthetic sequence contains an *Xba*I restriction site, positive clones containing insertions were initially identified by the presence of an additional *Xba*I site and were subsequently verified by sequencing (18).

Preparation of yeast total RNA and Northern blot analysis. The plasmids were introduced into strain B-7467 by the lithium acetate method (12). Yeast cells were grown in synthetic complete medium minus uracil (SC-Ura medium) with dextrose substituted with raffinose. Approximately 40 ml of yeast cells was harvested at about 2×10^7 to 3×10^7 cells per ml, washed in 1 ml of cell washing buffer (0.1 M LiCl, 0.1 M Tris-HCl [pH 7.5], 0.1 mM EDTA), and resuspended

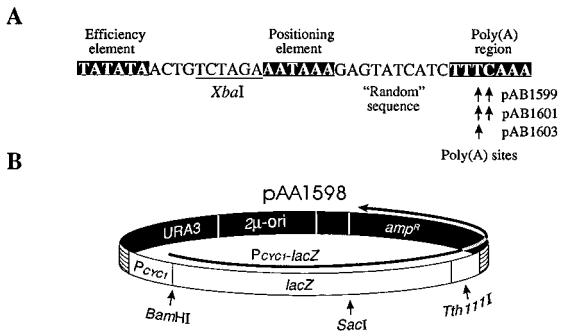
* Corresponding author. Mailing address: Department of Biochemistry, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642. Phone: (716) 275-2766. Fax: (716) 271-2683. Electronic mail address: fsrm@bphvax.biophysics.rochester.edu.

† Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125.

TABLE 1. Oligonucleotides used for PCR and Northern analysis and for construction of synthetic signals

Oligonucleotide	Sequence
OL85-86	TCTATAGACACGCAAACAC
OL95-43	GACGAGCTCGGATCCTGCAGTTTTTT
OL95-83	CACCAGCAGCAGTTTTTCCAG
OL95-85	CTTTGAAAGATGATACTCTTTATTTCTAGACAGTTATATA
OL95-86	GTTTGAAAGATGATACTCTTTATTTCTAGACAGTTATATA
OL95-87	GTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAA
OL95-88	GATCTTTGAAAGATGATACTCTTTATTTCTAGACAGTTATATA
OL95-89	GATCTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAA
OL95-90	CTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAA
OL95-91	CTTTGAAAGATGATACTCTTTATTTCTAGACAGTTATATAGAGCT
OL95-92	CTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAAGAGCT
OL95-99	TTGCACCACAGATGAAACGCCGAGTTAACGCCATC
OL95-100	CCTCTGACACATGCAGCTC

in 0.5 ml of yeast lysis buffer (0.1 M LiCl, 0.1 M Tris-HCl [pH 7.5], 0.1 mM EDTA, 0.5% sodium dodecyl sulfate [SDS]). An equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) and 0.5 g of glass beads (0.45-mm diameter) were added, and the cells were broken by vortexing four times for 15 s. The tubes were kept on ice between rounds of vortexing. The lysate was centrifuged at 10,000 × *g* for 10 min, and the supernatant was extracted twice with an equal volume of phenol-chloroform-isoamyl alcohol. RNA was precipitated with ethanol and dissolved in 40 μl of sterile distilled water. Twenty micrograms of yeast total RNA was separated by electrophoresis through a 1.2% formaldehyde agarose gel with 0.1 μg of ethidium bromide per ml by the method described by Fourney et al. (4). The fractionated RNA was transferred to a nitrocellulose filter. After 2 h of baking at 80°C in a vacuum oven, the filter was prehybridized at 42°C overnight in hybridization solution (50% deionized formamide, 50 mM NaH_2PO_4 [pH 6.5], 5× SSC, 1× Denhardt's solution, 0.25 mg of sheared and denatured calf thymus DNA per ml [20]). A ^{32}P-end-labeled probe, OL95-99, was then added to the hybridization solution to a final concentration of 5×10^5 cpm/ml. Hybridization was performed at room temperature for 16 h. The filter was then washed with 1× SSC–0.1% SDS twice for 20 min at room temperature and once for 10 min at 42°C. Autoradiography was performed at −70°C with an intensifying screen.



C

Restriction site	Inserted oligonucleotide duplex	Orientation	Plasmid	Strain
*Bam*HI	5′ GATCTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAA ATATATTGACAGATCTTTATTTCTCATAGTAGAAAGTTTCTAG 5′	+	pAB1599	B-9740
		−	pAB1600	B-9741
*Sac*I	CTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAAGAGCT 3′ 3′ TCGAGATATATTGACAGATCTTTATTTCTCATAGTAGAAAGTTTC	+	pAB1601	B-9742
		−	pAB1602	B-9743
*Tth*111I	5′ CTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAA ATATATTGACAGATCTTTATTTCTCATAGTAGAAAGTTTG 5′	+	pAB1603	B-9744
	5′ GTATATAACTGTCTAGAAATAAAGAGTATCATCTTTCAAA ATATATTGACAGATCTTTATTTCTCATAGTAGAAAGTTTC 5′	−	pAB1604	B-9745

FIG. 1. (A) The sequence of the synthetic 3′-end-forming signal. The efficiency element, positioning element, and poly(A) region are shown in solid boxes and are indicated above the sequences. A 10-bp random sequence that separates the poly(A) region and the positioning element is denoted; an *Xba*I restriction site which, along with a 4-bp random sequence, separates the efficiency and positioning elements is also denoted. The actual poly(A) sites determined by rapid amplification of cDNA ends PCR for the transcripts from pAB1599, pAB1601, and pAB1603 are indicated by arrows. (B) Schematic representation of the parental plasmid pAA1598 (pLG-312) (15). The *lacZ* open reading frame is fused to the *CYC1* promoter, which is denoted by P_{CYC1}. The longest transcript that starts from the *CYC1* promoter and terminates at the distal end of the Amp^r (β-lactamase) gene is shown. The positions of the three unique restriction sites, at which the synthetic signal was inserted, are denoted. (C) Details of the duplex oligonucleotides that were inserted in the parental plasmid pAA1598. One duplex oligonucleotide was inserted into the *Bam*HI and *Sac*I sites with two possible orientations: +, the orientation with which the transcripts of the fusion gene contain the sequence shown in panel A, and −, the opposite orientation. In contrast, to insert the synthetic sequence into the *Tth*111I site with two orientations, two different duplex oligonucleotides were used. The plasmid and strain numbers are shown on the right.

PCR mapping of the 3′ ends of *CYC1-lacZ* transcripts. PCR mapping of 3′ mRNA termini was conducted by the method described by Russo et al. (16, 17) except that after reverse transcription, one round of PCR amplification, consisting of 35 cycles, was performed with following primers: OL95-43 and OL85-86 for the transcripts with poly(A) sites adjacent to the *Bam*HI site, OL95-43 and OL95-83 for those with poly(A) sites adjacent to the *Sac*I site, and OL95-43 and OL95-100 for those with poly(A) sites adjacent to the *Tth*111I site. Note that OL95-43 is identical to the primer 2 that we previously used (16). The PCR products were separated on a 6% polyacrylamide–urea sequencing gel along with a sequencing ladder, which was the product of a sequencing reaction using the control DNA M13mp18 and −40 primer included in the Sequenase version 2.0 kit (United States Biochemicals).

Computer analysis. To obtain a 10-bp random sequence, a TTTAAAGGCC sequence was entered and randomized by using the University of Wisconsin Genetics Computer Group SHUFFLE software (version 7.0). The yeast sequences that are homologous to the OL95-83 sequence were searched for by using the University of Wisconsin Genetics Computer Group FASTA software.

RESULTS

The synthetic signal is sufficient to direct mRNA 3′-end formation in a *CYC1-lacZ* fusion gene. On the basis of our previous studies (6, 7, 16), we reasoned that three elements were required to form an efficient poly(A) site in yeast cells: the efficiency element, the positioning element, and a poly(A) region. The optimum sequences for efficiency and positioning elements are TATATA and AATAAA (or AAAAAA), respectively (7, 10). Several lines of evidence indicated that polyadenylation preferentially occurred at $Py(A)_n$ sequences and that the *CYC1* poly(A) site TTTCAAA appeared to be optimum (8, 16). In addition, the spacing between the elements also affected the efficiency of mRNA 3′-end formation (16). The positioning element usually produces poly(A) sites 10 to 20 bp downstream. Although the efficiency element can be located hundreds of base pairs upstream of the actual poly(A) site, a 10- to 20-bp spacing between the efficiency and positioning elements seems to be preferred (16). Therefore, we synthesized an oligonucleotide, shown in Fig. 1A, which contains the efficiency element TATATA, the positioning element AATAAA, and the *CYC1* poly(A) region TTTCAAA, as well as two 10-bp random spacing sequences. To avoid unusual

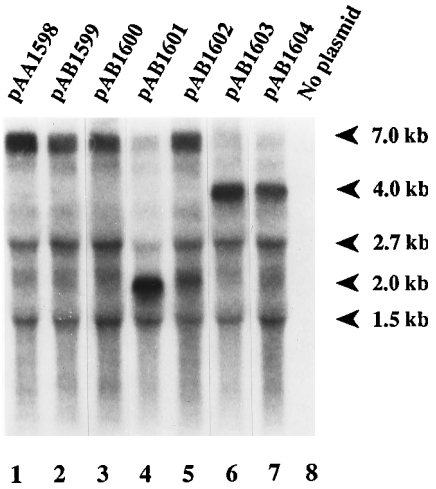


FIG. 2. Northern blot analysis of the transcripts from the *CYC1-lacZ* fusion genes with or without insertions. Lane 1, the host cells containing the parental plasmid pAA1598, which produced heterogeneous *CYC1-lacZ* transcripts; lanes 2 to 7, the host cells transformed with the constructs that contained the synthetic sequence insertions with two orientations at the *Bam*HI, *Sac*I, and *Tth*111I restriction sites (Fig. 1C); lane 8, the host cells only. The various transcripts, indicated on the right, include the following: 1.5-, 2.7-, and 7-kb transcripts produced by the parental plasmid pAA1598; the 2-kb transcripts produced by the plasmid pAB1601; and the 4-kb transcript produced by the plasmids pAB1603 and pAB1604. Note that these 2- and 4-kb transcripts were absent in pAA1598.

sequences, such as those containing runs of the same nucleotides, or sequences with extremely high or extremely low A+T content, we entered and shuffled a sequence with an A+T content of 60% (see Materials and Methods) and chose two relatively balanced sequences as spacing sequences (Fig. 1A). To facilitate the detection of insertions, an *Xba*I restriction site was included in the spacing sequence between the efficiency element and the positioning element.

The synthetic poly(A) signal was introduced at three sites in the transcription unit of a *CYC1-lacZ* fusion gene (Fig. 1B) (5), which was composed of a *CYC1* promoter, the *CYC1* 5′ untranslated region, and the *lacZ* open reading frame. Because of the lack of a strong poly(A) signal, the transcripts from this fusion gene are heterogeneous (15). The longest transcript is approximately 7 kb long with endpoints at the distal end of the Amp^r (β-lactamase) gene (Fig. 1B). The duplex oligonucleotides (Fig. 1C) containing the synthetic 3′-end-forming signal were inserted with two possible orientations into three unique restriction sites, *Bam*HI, *Sac*I, and *Tth*111I, within the transcription unit, which were located, respectively, 0.1, 2.0, and 4.0 kb downstream of the *CYC1* major transcription initiation site (14). The plasmids were introduced into the yeast host B-7467, and Northern blot analysis was performed. As expected, the *CYC1-lacZ* fusion gene produced several major transcripts of 1.5, 2.7, and 7 kb, as well as a smeared background (Fig. 2; compare lanes 1 and 8). In addition to these transcripts, a 2-kb transcript appeared in the strain B-9742 harboring pAB1601 (Fig. 2, lane 4), indicating that the synthetic signal inserted at the *Sac*I site caused 3′-end formation even in the translated region. This insertion resulted in 86 and 55% reductions, respectively, in the levels of 7-kb transcript and 2.7-kb transcript, if the 1.5-kb transcript is considered an internal control. The synthetic signal was inserted in an inverted orientation in pAB1602, in which a transcript of approximately 2.1 kb was slightly enhanced (Fig. 2, lane 5). It should be noted that insertion in an inverted orientation destroys the positioning element and the poly(A) site sequence but still retains the efficiency element TATATA. The faint 2.1-kb band could represent a cryptic poly(A) site that was enhanced by the efficiency element TATATA. In pAB1603, the signal was inserted in the *Tth*111I site located downstream of the *lacZ* translated region. At this position, the synthetic signal directed the formation of a 4-kb transcript and almost eliminated the 7-kb transcript, resulting in more than a 90% diminution (Fig. 2, lane 6). Insertion of the signal in an inverted orientation at the *Tth*111I site (pAB1604) also resulted in the appearance of a 4-kb transcript, but at a lower relative level than for pAB1603 (Fig. 2, lane 7). As expected, this insertion (pAB1604) resulted in approximately 90% reduction in the level of the 7-kb transcript. The 4-kb transcript in pAB1604 could be generated by the TATATA sequence of the inverted synthetic sequence, combined with some cryptic signals downstream in the *lacZ* sequence. If this is the case, the poly(A) sites of this transcript should be downstream from the inserted sequence. The 3′-end-forming signal was also inserted into the *Bam*HI site (pAB1599 and pAB1600), where the *lacZ* open reading frame was fused to the *CYC1* promoter and 5′ untranslated region. Since the probe (OL95-99) used for the Northern blot analysis was a segment of *lacZ* sequence approximately 400 bp downstream from the *Bam*HI site, any transcripts produced by the poly(A) signal inserted at this site could not be detected. In addition, such transcripts should be less than 150 nt long, and short transcripts are generally difficult to detect by agarose gel Northern blot analysis.

The 3′ ends of truncated transcripts were located in the poly(A) region of the synthetic signal. In order to detect the short transcripts produced by pAB1599 and pAB1600 and to precisely map the 3′ termini of the transcripts formed by the synthetic poly(A) signal at the other two restriction sites, PCR 3′-end mapping analysis was carried out. As shown in Fig. 3A, pAB1599 did produce a major transcript, whose 3′ ends mapped to the last T residue and the C residue in the TTTCAAA sequence of the synthetic signal insert. Furthermore, pAA1598 and pAB1600 lacked this transcript, indicating that it is the result of the insertion at the *Bam*HI site. Several minor transcripts with 3′ ends further downstream and another minor transcript, which were not seen in the parent plasmid pAA1598, were detected in pAB1599 and pAB1600, respectively (Fig. 3A). We interpret this as indicating the presence of cryptic poly(A) sites, which were enhanced by the efficiency element TATATA that existed in both the synthetic sequence and the inverted synthetic sequence.

Similarly, the 3′ ends of the 2-kb transcript produced in pAB1601 were also mapped to the last T residue and the C residue in the TTTCAAA sequence of the inserted signal (Fig. 3B, lane 5). In addition to this specific PCR product, several nonspecific longer products were detected in pAB1601, pAB1602, and the parental plasmid pAA1598. Because there was a 44-bp insert in pAB1601 and pAB1602, these bands could not be produced by cryptic poly(A) sites in the plasmids. Otherwise, the bands in the parental plasmid pAA1598 should be 44 nt shorter than those in pAB1601 and pAB1602. We propose that they may be PCR artifacts that came from the transcripts of the yeast host. In this regard, we found that the oligonucleotide OL95-83, the upstream primer used for rapid amplification of cDNA ends-PCR experiments, was highly homologous to a sequence located approximately 40 bp downstream of the translation termination codon of a putative yeast open reading frame (GenBank locus YSCCHR1RAA). A total of 14 of 21 nt in OL95-83 and this 3′ untranslated region sequence, including those at the 3′ end of the primer, were identical. It is likely that the bands seen in all three plasmids (Fig. 3B) represented this putative transcript amplified by the PCR primers OL95-83 and

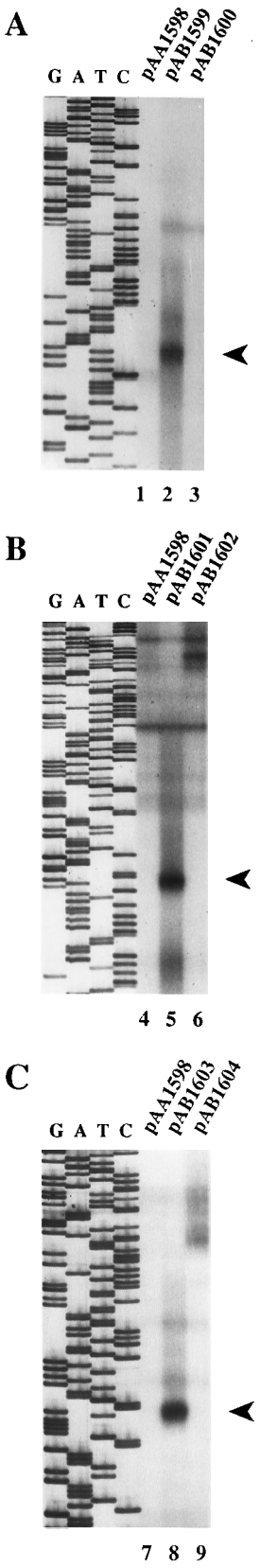


FIG. 3. PCR 3′-end mapping of the *CYC1-lacZ* transcripts from the parental plasmid pAA1598 and from the plasmids containing inserts. The 3′ ends of transcripts that were adjacent to the *Bam*HI, *Sac*I, and *Tth*111I restriction sites were mapped, respectively, by three rapid amplification of cDNA ends PCR experiments (A to C). The ^{32}P-radiolabeled 3′-end mapping products were separated by electrophoresis through a 6% polyacrylamide–urea sequencing gel, along with a sequencing ladder, which was the product of a sequencing reaction using the control DNA M13mp18 and −40 primer included in the Sequenase version 2.0 kit (United States Biochemicals). By using the sequencing ladder as size standards, the length of the PCR products could be estimated with an accuracy of ±1 bp, allowing the determination of the approximate positions of the major poly(A) sites as presented in Fig. 1A. The major poly(A) sites for the transcripts from pAB1599 (lane 2), pAB1601 (lane 5), and pAB1603 (lane 8) are indicated by arrows. A shorter exposure of the gels showed that the major bands in lanes 2 and 5 were doublets while the major band in lane 8 was a single band.

OL95-43, because of this unexpected homology. In addition to these nonspecific bands, it should be noted that there was a relatively strong band in pAB1602 that was not found in pAA1598 and pAB1601 (Fig. 3B, lane 6). We suggest that this band represented a poly(A) site created by the TATATA element of the inverted synthetic sequence in pAB1602 and that this poly(A) site resulted in the 2.1-kb transcript detected in the Northern blot analysis (Fig. 2, lane 5). Finally, the 3′ ends of the transcripts with poly(A) sites adjacent to the *Tth*111I site were mapped by rapid amplification of cDNA ends-PCR. The 4-kb transcript of pAB1603 had a unique major 3′ end at the last T of the TTTCAAA sequence (Fig. 3C, lane 8), as well as several minor sites further downstream. As expected, the 3′ termini of the 4-kb transcript from pAB1604 were mapped to a sequence approximately 60 nt downstream from the TATATA sequence contained in the inverted synthetic signal (Fig. 3C, lane 9). Because of the lack of a strong positioning element and an optimal poly(A) site sequence, these poly(A) sites were scattered over a large region.

DISCUSSION

Although the efficiency element, the positioning element, and the poly(A) site sequence have been extensively studied with a cauliflower mosaic virus sequence (10), the *CYC1* system (7), and the yeast *ADH1* gene (8), it is still unclear if there are any other distinct elements involved in mRNA 3′-end formation. In this study we have tested whether these three elements constitute a functional poly(A) signal in yeast cells by themselves. The results showed that a synthetic sequence containing the optimum sequences of the three elements directed mRNA 3′-end formation in vivo in a *CYC1-lacZ* fusion gene, indicating that the elements are sufficient to specify a poly(A) site. Because the spacing sequences were generated by a computer program and were not identical or similar to any 3′-end-forming signals detected in *S. cerevisiae*, we believe that they could not significantly contribute to the formation of mRNA 3′ ends except just by separating the three elements. Our findings provide further support for our previous proposal that *S. cerevisiae* employs three degenerate elements to direct 3′-end formation of mRNA (7).

In this study, the synthetic 3′-end-forming signal was inserted at three different restriction sites, apparently containing different sequence contexts. Considering that the two 10-bp spacing sequences inside the synthetic signal have an A+T content of 60% and that the 100-bp regions adjacent to both the *Sac*I and *Tth*111I restriction sites have an A+T content of approximately 40% (13, 19), the overall content of A and T residues in the regions of 3′-end formation was not as high in *CYC1-lacZ* as it is in the 3′ regions of most yeast genes, e.g., *CYC1*. Thus, an A+T-rich context is not absolutely required for 3′-end formation of yeast mRNAs. However, the bias for A+T-rich sequences in the 3′ untranslated regions, as well as in 5′ untranslated regions, is generally observed in yeast genes.

Such A+T-rich sequences may simply reflect the general need to prevent secondary structures that inhibit any one of a number of processes.

Consistent with the appearance of the 2-kb transcript in pAB1601 and the 4-kb transcript in pAB1603 and pAB1604, the levels of 7- and 2.7-kb transcripts were greatly reduced by the insertions, suggesting that a strong 3′-end formation event occurring upstream can affect the 3′-end formation downstream. However, in contrast to the synthetic signal, the 83-bp *CYC1* 3′ sequence, which was inserted at the *Sac*I and *Tth*111I sites, completely eliminated the 2.7- and 7-kb transcripts of the *CYC1-lacZ* fusion gene (15). This result may indicate that the 83-bp *CYC1* sequence is more efficient than the synthetic signal. In addition to the well-defined *CYC1* 3′-end-forming signals (6, 16), the 83-bp *CYC1* region contains highly A+T-rich flanking sequences that may alleviate possible adverse effects of the *lacZ* sequence, such as secondary structures. In this regard, it will be interesting to see if any highly A+T-rich flanking sequences can provide a better context for the synthetic signal.

Although the synthetic signal is sufficient to direct 3′-end formation at the three positions along the *CYC1-lacZ* fusion gene, the patterns of the poly(A) sites at the three positions were not identical. When the synthetic signal was inserted into the *Bam*HI and *Sac*I sites, the poly(A) sites were mapped to the last T residue and the C residue of the TTTCAAA sequence. In pAB1603, which had the insertion at the *Tth*111I site, the poly(A) site was mapped to the last T residue of the TTTCAAA sequence. It should be mentioned here that in the wild-type *CYC1* gene, the C residue is the nucleotide to which the poly(A) tail is added. The slight difference in the selection of poly(A) sites could be due to the adjacent sequences, which may affect the precise selection of the poly(A) site (7).

The formation of mRNA 3′ termini is an important step in gene expression. It is now clear that the capacity of the 3′-end-forming signal can greatly affect the level of gene expression. In this study, creation of a poly(A) signal in the 3′ untranslated region of the fusion gene in the pAB1603 plasmid resulted in a severalfold increase in the level of full-length transcripts (Fig. 2, lane 6) and presumably increased the production of the protein. The finding that three simple sequences can constitute a poly(A) signal may be helpful for expressing high levels of gene products, especially those products that are not of yeast origin.

ACKNOWLEDGMENTS

We are grateful to L. Guarente (MIT) for providing the plasmid pLG-312. We thank E. M. Phizicky and J. S. Butler (University of Rochester) for helpful discussions. We also thank L. Comfort for synthesis of the oligonucleotides used in this study.

This work was supported by National Institutes of Health research grant RO1 GM12702.

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. **Butler, J. S., and T. Platt.** 1988. RNA processing generates the mature 3′ end of yeast *CYC1* messenger RNA *in vitro*. Science **242:**1270–1274.
3. **Egli, C. M., C. Springer, and G. H. Braus.** 1995. A complex unidirectional signal element mediates *GCN4* mRNA 3′ end formation in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **15:**2466–2473.
4. **Fourney, R. M., J. Miyakoshi, R. S. Day III, and M. C. Paterson.** 1988. Northern blotting: efficient RNA staining and transfer. Focus **10:**5–7.
5. **Guarente, L., and T. Mason.** 1983. Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. Cell **32:**1279–1286.
6. **Guo, Z., P. Russo, D.-F. Yun, J. S. Butler, and F. Sherman.** 1995. Redundant 3′ end-forming signals for the yeast *CYC1* mRNA. Proc. Natl. Acad. Sci. USA **92:**4211–4214.
7. **Guo, Z., and F. Sherman.** 1995. 3′-end-forming signals of yeast mRNA. Mol. Cell. Biol. **15:**5983–5990.
8. **Heidmann, S., B. Obermaier, K. Vogel, and H. Domdey.** 1992. Identification of pre-mRNA polyadenylation sites in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12:**4215–4229.
9. **Henikoff, S., J. D. Kelley, and E. H. Cohen.** 1983. Transcription terminates in yeast distal to a control sequence. Cell **33:**607–614.
10. **Irniger, S., and G. H. Braus.** 1994. Saturation mutagenesis of a polyadenylation signal reveals a hexanucleotide element essential for mRNA 3′ end formation in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **91:**257–261.
11. **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.
12. **Ito, H., Y. Fukuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153:**163–168.
13. **Kalnins, A., K. Otto, U. Ruther, and B. Muller-Hill.** 1983. Sequence of the *lacZ* gene of *Escherichia coli*. EMBO J. **2:**593–597.
14. **Li, W.-Z., and F. Sherman.** 1991. Two types of TATA elements for the *CYC1* gene of the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. **11:**666–676.
15. **Osborne, B. I., and L. Guarente.** 1988. Transcription by RNA polymerase II induces changes of DNA topology in yeast. Genes Dev. **2:**766–772.
16. **Russo, P., W.-Z. Li, Z. Guo, and F. Sherman.** 1993. Signals that produce 3′ termini in *CYC1* mRNA of the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. **13:**7836–7849.
17. **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.
18. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
19. **Sutcliffe, J. G.** 1979. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43:**77–90.
20. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **81:**5201–5205.
21. **Wahle, E., and W. Keller.** 1992. The biochemistry of 3′-end cleavage and polyadenylation of messenger RNA precursors. Annu. Rev. Biochem. **61:** 419–440.
22. **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.
23. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28:**563–573.