MOLECULAR AND CELLULAR BIOLOGY, Apr. 1991, p. 2004–2012
0270-7306/91/042004-09$02.00/0

Vol. 11, No. 4

Point Mutations Upstream of the Yeast *ADH2* Poly(A) Site Significantly Reduce the Efficiency of 3′-End Formation

LINDA E. HYMAN, STEPHANIE H. SEILER,† JOHN WHORISKEY, AND CLAIRE L. MOORE*
Department of Molecular Biology and Microbiology, Tufts University, Boston, Massachusetts 02111-1800

Received 10 December 1990/Accepted 11 January 1991

The sequences directing formation of mRNA 3′ ends in *Saccharomyces cerevisiae* are not well defined. This is in contrast to the situation in higher eukaryotes in which the sequence AAUAAA is known to be crucial to proper 3′-end formation. The AAUAAA hexanucleotide is found upstream of the poly(A) site in some but not all yeast genes. One of these is the gene coding for alcohol dehydrogenase, *ADH2*. Deletion or a double point mutation of the AAUAAA has only a small effect on the efficiency of the reaction, and in contrast to the mammalian system, it is most likely not operating as a major processing signal in the yeast cell. However, we isolated point mutations which reveal that a region located approximately 80 nucleotides upstream of the poly(A) site plays a critical role in either transcription termination, polyadenylation, or both. These mutations represent the first point mutations in yeasts which significantly reduce the efficiency of 3′-end formation.

The formation of the 3′ ends of eukaryotic mRNAs involves several steps, including termination of transcription, cleavage of precursor RNA, and addition of poly(A). In higher eukaryotes, it is known that termination of transcription by RNA polymerase II occurs downstream from the end of the mature mRNA. The actual end of the mRNA is produced by cleavage of the primary transcript and the subsequent addition of about 200 adenylate residues (for reviews, see references 30 and 36). The relationship between termination of transcription and RNA processing is not well understood, although it appears that termination requires valid RNA-processing sites (13, 17, 48). Polyadenylation has been studied both in vivo and in vitro in mammalian cells. From these studies, it is clear that sequences on the precursor RNA as well as several factors, including specificity factors, cleavage factors, and a poly(A) polymerase, are required for proper 3′-end formation (10, 11, 18, 31, 44). The *cis*-acting sequence which is most highly conserved in metazoans is the hexanucleotide AATAAA, usually found within 30 nucleotides (nt) 5′ of the polyadenylation site. This sequence is clearly required for assembly of the processing complex and for cleavage and polyadenylation of the mRNA precursor (30, 36). In contrast, the lack of conservation of this sequence is apparent in a survey of gene sequences from the yeast *Saccharomyces cerevisiae* as well as higher plant species (26). In yeasts, this sequence appears at the 3′ end of mRNA in approximately 50% of yeast genes (24a). Several other consensus sequences have been proposed for yeast 3′-end formation (23, 51), but none are as conserved as the AATAAA signal found in metazoans. Therefore, although all yeast mRNAs are polyadenylated and a variety of precursors are accurately cleaved in yeast extracts (1, 7, 8), the signals directing these events are not clear.

The relationship between transcription termination and RNA processing in yeasts is also undefined. Recent work from several groups has suggested that transcription termination occurs close to the end of the mature transcript. These data are based on measuring plasmid destabilization which results from transcription through a centromere (40) or an autonomously replicating sequence (42). For example, when a sequence from the 3′ end of the yeast *CYC1* gene is placed between a promoter and the centromere, the plasmid is stabilized, suggesting that transcription termination is signalled by this *CYC1* sequence (40). Nuclear run-on experiments support this hypothesis (33). However, these experiments do not address the relationship between the processing of the transcripts and transcription termination, and this issue remains elusive.

In this report, we describe an in vivo assay which uses β-galactosidase production to quantitate the efficiency of 3′-end formation. We show that the AAUAAA sequence important for 3′-end formation of mammalian mRNAs is not required for the processing of a yeast transcript which naturally contains this sequence. We found instead that sequences further upstream of the poly(A) site are crucial for transcription termination and/or RNA processing. In addition, we describe the first point mutations which significantly reduce the efficiency of 3′-end formation in yeasts.

MATERIALS AND METHODS

Reagents. All restriction enzymes, modifying enzymes, DNA size standards, and linkers were purchased from New England BioLabs and were used according to the manufacturer's instructions. RNA polymerases were purchased from Promega and Stratagene. Nonradioactive nucleotides and the pT7T3 vector were from Pharmacia, Inc. Radiolabeled nucleotides were from Dupont, NEN Research Products. T2 RNase was from Bethesda Research Laboratories. Sequenase was purchased from U.S. Biochemical Corp.

Plasmid constructions. The plasmids pL101 and PL102 were constructed as follows. A 327-bp fragment from the 3′ end of the *ADH2* gene was subcloned from the plasmid pYEPADH2(E/B) (38) into the *Hin*cII-*Sph*I site of pT7T3 18. The *ADH2* sequence was inserted into the yeast expression vector pHZ18Δ2 Sma (Fig. 1). This plasmid contains the yeast 2μm origin of replication and the *URA3* gene for selection, in addition to the pBR322 origin and the ampicillin-selectable marker for propagation in bacteria. pHZ18Δ2 Sma was created by converting the unique *Sal*I site present in pHZ18Δ2 (34, 45) to an *Sma*I site by ligation of *Sma*I

* Corresponding author.
† Present address: Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY 10461.

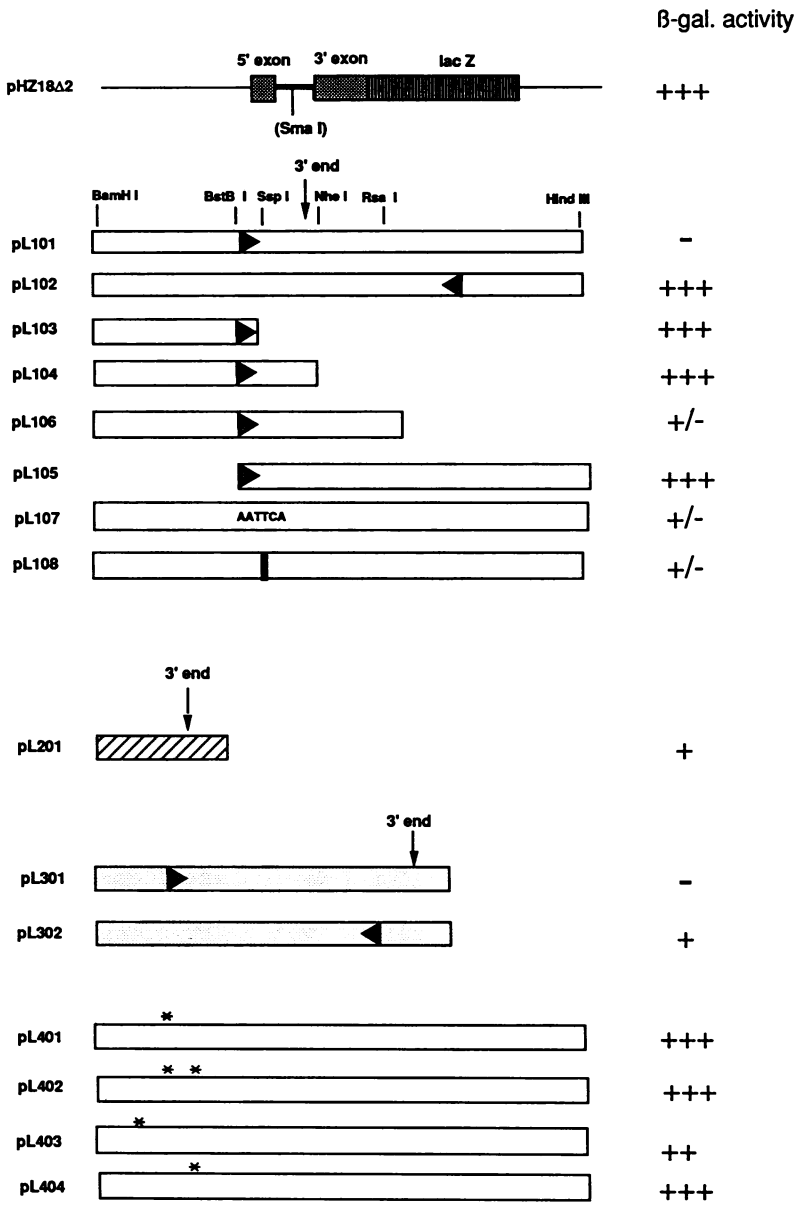

FIG. 1. Map of fusion genes used to define sequences important for 3'-end formation and their corresponding β-galactosidase activities. pHZ18Δ2 is a yeast 2μm vector containing *URA3* as a selectable marker and an rp51-*lacZ* fusion gene which expresses relatively high levels of β-galactosidase. The plasmids pL101 to pL108 contain sequences from the 3' end of the *ADH2* gene (open boxes), pL201 contains 82 nt from the 3' end of *CYC1* (striped box), pL301 and pL302 contain 310 nt from the 3' end of the *ACP2* gene (stippled boxes), and pL401 to pL404 contain mutations in the *ADH2* sequence which affect 3'-end formation. The shaded triangle depicts the site and orientation of the AATAAA sequence. Arrows point to 3' end of fusion genes, and * depicts sites of point mutations. The β-galactosidase activities are expressed relative to that of pHZ18Δ2. The activity measurements are based on the plate color assay, where +++ is very dark blue and − is white.

linkers into the filled-in *Sal*I site (3). All subsequent cloning into pHZ18Δ2 Sma was performed by filling in of 5' overhangs generated by restriction digests and blunt-end ligation into the *Sma*I site. pL103 was generated by cleavage of pT7T3 ADH2 with *Bam*HI and *Ssp*I, filling in the 5' overhangs with the large fragment of DNA polymerase I, gel purifying the fragment away from the vector, and ligation to *Sma*I-cut pHZ18Δ2 Sma (3). pL104, pL105, and pL106 were constructed in a similar manner except that the starting plasmid was digested with *Nhe*I and *Bam*HI for pL104, with *Bst*BI and *Hin*dIII for pL105, and with *Bam*HI and *Rsa*I for pL106. The plasmid pL201 contains the 82-nt fragment from the 3' end of the *CYC1* gene and approximately 30 nt derived from the polylinker of pGEM. The plasmid pSB7 (described in reference 37) was digested with *Xho*I and *Sal*I and religated. The resulting plasmid containing the *CYC1* sequence was digested with *Hin*dIII and the ends were filled in, and then it was digested with *Sma*I and ligated into pHZ18Δ2 Sma. The plasmids pL301 and pL302 were derived from digestion of pWH7, a plasmid containing the *ACP2* gene (a gift from D. Kolodrubetz). pWH7 was digested with *Bam*HI and *Eco*RI to generate a 310-nt fragment which contained the 3' end of the *ACP2* gene (20). This was subcloned into pT7T3 to create pT7T3 ACP2. This plasmid was digested with

*Eco*RI and *Hin*dIII and, after the ends were filled in, cloned into the *Sma*I site of pHZ18Δ2 Sma. The plasmids pL107 and pL108 were constructed by oligonucleotide-directed site-specific mutagenesis on double-stranded DNA by the method of Inouye and Inouye (25). The mutagenesis was done on the pT7T3 ADH2 plasmid and the mutated ADH2 DNA was then subcloned into pHZ18Δ2 Sma by using the same strategy as for pL101. Mutations were verified by DNA sequencing.

RNA analysis. Total RNA was prepared as described by Ausubel et al. (3). RNase protection analysis was done as described by Hart et al. (22). Radiolabeled size standards were made by filling in 5′ overhangs from *Msp*II-cut pBR322 DNA. Northern (RNA) analysis was performed after electrophoresis on 1.5% agarose-formaldehyde gels (3). Radiolabeled *ADH2* probes for hybridization were made by digesting pT7T3 ADH2 with *Eco*RI and using this DNA as a template for T3 RNA polymerase. The *URA3* probe was made by digestion of pNSU114 (a gift from Neil Sugawara) with *Eco*RI, and this was used as a template for SP6 RNA polymerase. Densitometry was performed with the Scanjet densitometry software for the MacIntosh computer and also with an LKB ultroScan XL densitometer. Several different exposures of Northern blots were analyzed. The RNA fold programs are from the Genetics Computer Group (15) and are based on the methods of Zucker and Stiegler (52).

Chemical mutagenesis. Chemical mutagenesis was done with hydroxlyamine hydrochloride (Sigma), using two different protocols. The first utilized 10 μg of CsCl-purified pL101 DNA, treated with a freshly made solution of 0.5 M hydroxylamine hydrochloride (pH 7.0) and incubated at 37°C for 20 h. The reaction was terminated by the addition of 0.1 M NaCl and 0.1-mg/ml bovine serum albumin and two ethanol precipitations prior to transformation. The second protocol (12) started with 30 μg of CsCl-purified pL101 in 75% ethylene glycol. The hydroxylamine solution (final concentrations of 0.05 M hydroxlyamine and 0.02 M NaP_2O_7) was added to the DNA and incubated at 70°C for 45 min. The reaction was terminated by the addition of acetone to 14% and NaCl to 0.14 M. The DNA was ethanol precipitated two times prior to transformation.

Analysis of mutants and DNA sequencing. Yeast colonies which appeared blue on indicator plates after hydroxylamine mutagenesis were chosen for further analysis. Colonies were grown on appropriate plates, and total DNA was prepared from cells taken directly from the plate (41). This DNA was transformed into *Escherichia coli* (MC1061), and the transformation mixture was reintroduced into competent yeast cells to confirm the original phenotype. DNA from all colonies which were blue after the second transformation was prepared and transformed into *E. coli* as described by Sherman et al. (41) except that the cells were taken directly from the plate and an extra chloroform-isoamyl alcohol extraction and ethanol precipitation were included prior to transformation. DNA from these bacterial transformants was prepared with a Qiagen midi DNA preparation kit and sequenced by the double-stranded dideoxy sequencing technique with Sequenase (U.S. Biochemical Corp.). In all cases, both strands were sequenced. It is noteworthy that the sequence we obtained for the wild-type *ADH2* sequence differs from the published sequence (38) at four sites (positions 25, 31, 34, and 49) and that in all cases there is a T-to-C transition.

Yeast strains and yeast methods. The yeast strain used throughout this work is DB745 (α *ade leu ura*). All yeast methods are described by Sherman et al. (41). Yeast transformation was done by the LiCl method. β-Galactosidase activities were measured by the cell permeabilizing assay described by Ausubel et al. (3).

RESULTS

To determine whether the AATAAA sequence found in some yeast genes is important in 3′-end formation and to ask what additional sequences may be required for this reaction, we constructed a plasmid which allows us to use β-galactosidase activity as a measure of 3′-end formation in vivo. This plasmid, pHZ18Δ2, is shown schematically in Fig. 1. The key component of this plasmid is a fusion gene which consists of part of the ribosomal protein (rp) 51 gene (5′ exon, intron, and part of 3′ exon) fused in frame to the *lacZ* gene (34, 45). The fusion gene is controlled by the *GAL* upstream activating sequence. Under inducing conditions, it was transcribed, the pre-mRNA was spliced, and β-galactosidase was produced (Fig. 1 and 2B).

We reasoned that insertion of a sequence which directs transcription termination or polyadenylation (or both) into the intron would prevent the completion of full-length transcripts and reduce or abolish β-galactosidase activity. To test this, we inserted the 3′ end of the *ADH2* gene into the intron of the rp51 gene to create pL101. The 3′ end of the *ADH2* mRNA has been mapped by S1 nuclease protection (38), and an AAUAAA motif is found 17 nt upstream of this site, in a position consistent with it possibly functioning as a polyadenylation signal. If the sequence from the *ADH2* gene directs either transcription termination or 3′-end processing (or both), no β-galactosidase will be produced. If the *ADH2* site does not direct 3′-end formation, transcription will proceed through the *lacZ* gene, the rp51 intron will be spliced, and the fusion transcript will direct synthesis of functional β-galactosidase. The levels of β-galactosidase can be determined by both a plate assay (blue versus white colonies on indicator plates) and/or a soluble assay. In all cases, cells were grown in media containing galactose as a carbon source (to activate the promoter) and lacking uracil (to maintain the plasmid). Transcription from the original plasmid, pHZ18Δ2, directed the synthesis of β-galactosidase (Fig. 1 and 2B). Unlike pHZ18Δ2, plasmid pL101, containing 327 nt derived from the 3′ untranslated region of *ADH2*, did not produce high levels of β-galactosidase. The *ADH2* fragment inserted in the opposite orientation (pL102) resulted in high levels of β-galactosidase expression, demonstrating that the signal directing 3′-end formation of this mRNA is orientation dependent. In addition, this result indicates that additional sequence present in the intron of these fusion plasmids does not interfere with splicing.

The absence of β-galactosidase from the strain containing the *ADH2* sequence in the sense orientation does not establish that the fusion transcript contains a 3′ end corresponding to the wild-type *ADH2* transcript. To address this question, we prepared RNA from cells harboring pL101 or pL102 or pYEPADH2(E/B), a plasmid containing the entire wild-type *ADH2* gene (38). The ends of RNAs containing the *ADH2* sequence were determined by RNase protection analysis. RNA derived from cells containing pL101 (Fig. 3A, lane 3) yielded two protected fragments, a major one of 144 nt and a minor one of 124 nt. The major poly(A) site mapped to a stretch of Ts at positions 133 to 137 (see Fig. 5A), and the minor site mapped to position 118. This is identical to the 3′ ends of the endogenous *ADH2* mRNA (Fig. 3A, lane 2). The difference in sizes of the protected fragments is due to non-*ADH2* sequences which are in the fusion transcript and

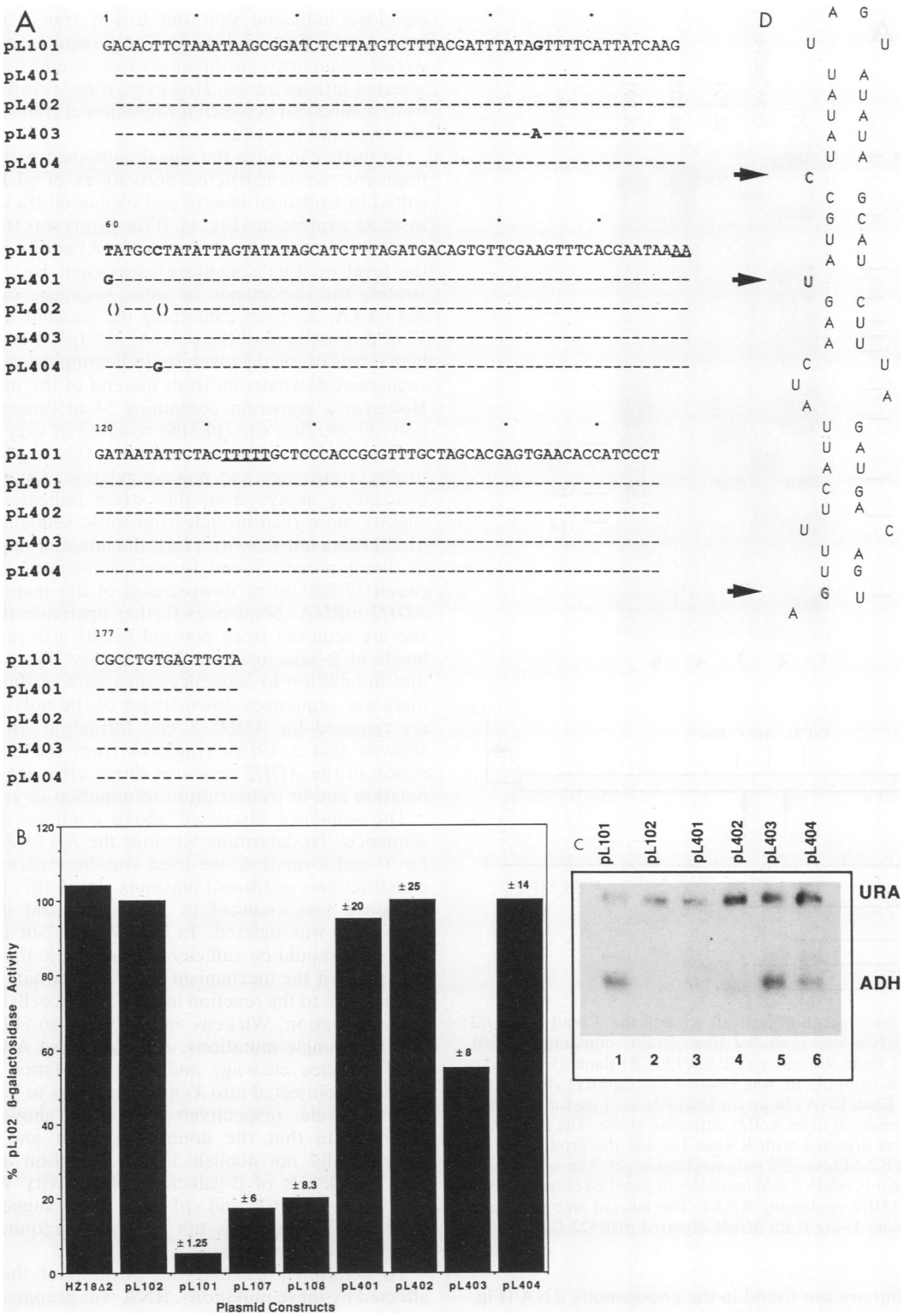

FIG. 2. Analysis of mutations induced by hydroxylamine. pL101 DNA was treated with hydroxylamine hydrochloride before transformation into yeast cells. Colonies which appeared blue on indicator plates were chosen for further analysis. Plasmid DNA from yeast cells was transformed into *E. coli*, and the DNA sequence was determined by conventional procedures. (A) Minimal sequence required for 3'-end formation. In addition, the sequence changes detected in the mutations pL401 to pL404 are shown. The major and minor polyadenylation sites are underlined. (B) β-Galactosidase activity of the mutants relative to that of the wild-type plasmid. β-Galactosidase activity was determined by analysis of solubilized cell extracts with *o*-nitrophenyl-β-D-galactopyranoside as the enzyme substrate. The data were compiled by averaging the results of approximately 10 independent experiments for each strain. (C) Northern blot analysis of the mutants compared with the wild-type strain. The blot was hybridized to a *URA3* probe and an *ADH2* probe. (D) Part of the predicted secondary structure of the *ADH2* sequence required for proper 3'-end formation. The arrows point to the positions of the point mutations.

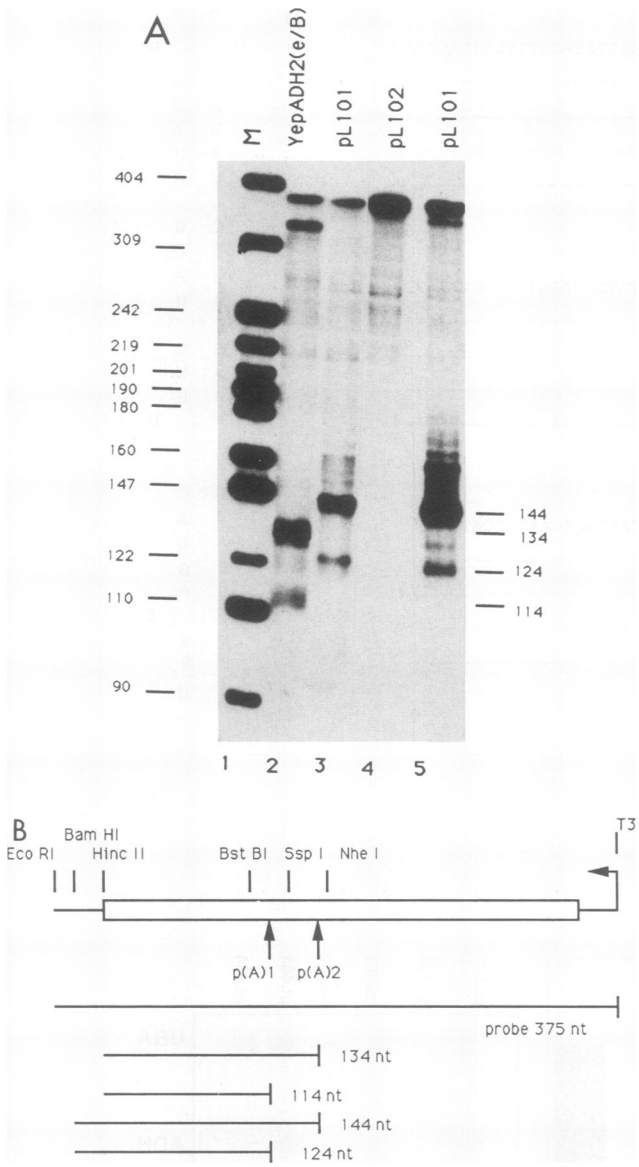

FIG. 3. RNase protection analysis to map the 3′ end of *ADH2* RNAs. Total RNA was prepared from strains containing pL101 (lane 3), pL102 (lane 4), and pYEPADH2(E/B) (lane 2). Poly(A)-containing RNA from pL101 was isolated by oligo(dT) chromatography (lane 5). Each RNA [10 μg for total RNAs, 1 μg for poly(A)$^+$ RNAs] was hybridized to an *ADH2* antisense probe. The hybridization mixture was digested with RNase T2, and the products were analyzed on an 8.3 M urea–8% polyacrylamide gel. The sizes of the protected fragments (shown schematically in panel B) correspond to the 3′ end of *ADH2*-containing RNAs. The labeled size standards (nucleotides) (lane 1) are from *Msp*II-digested pBR322 DNA.

in the probe but are not found in the endogenous RNA (Fig. 3B) and does not reflect a difference in the cleavage sites. The high-molecular-weight band running at about position 350 most likely represents incomplete digestion of the probe, since this species is observed in pL102 (lane 4), which does not contain any RNA complementary to the probe. The endogenous *ADH2* RNA was transcribed only when the cells were grown in the presence of ethanol and was not detectable in cells grown with galactose as the sole carbon source (Fig. 3A, lane 4). Analysis of RNA selected on oligo(dT)-cellulose indicated that the fusion transcript contains a poly(A) tail (Fig. 3A, lane 5). These results establish that the *ADH2* sequence can direct correct 3′-end formation when inserted into an intron. However, these results do not allow us to distinguish between termination of transcription and/or cleavage and polyadenylation.

To further localize the signal contained within this 327-nt fragment, we constructed derivatives of pL101 which resulted in truncated inserts and evaluated these for β-galactosidase expression (Fig. 1). If the insert was truncated at the *Ssp*I site, 5 nt upstream of the poly(A) addition site (pL103), the level of β-galactosidase expression was high, demonstrating the importance of some sequence downstream of AATAAA. A strain containing the plasmid which includes 17 nt downstream of the poly(A) site (pL104) also produced high levels of β-galactosidase, indicating the requirement for sequences downstream from the end of the mature mRNA. However a construct containing 54 nt downstream of the poly(A) addition site (pL106) resulted in very low levels of β-galactosidase, demonstrating that this region is sufficient to direct cleavage and polyadenylation. Quantitative β-galactosidase assays from this strain indicate that there is slightly more readthrough from these sequences than from pL101 (data not shown). Thus, the minimal sequence needed to direct proper 3′-end formation includes sequences between 17 and 54 nt downstream of the mature end of the *ADH2* mRNA. Sequences further upstream than the *Bst*BI site are required since plasmid pL105 also resulted in high levels of β-galactosidase. In summary, these results show that in addition to sequences upstream of the poly(A) site, there are sequences downstream of the poly(A) site which are required for *ADH2* 3′-end formation. In addition, we showed that a 191-nt fragment from the 3′ untranslated region of the *ADH2* gene can direct cleavage and polyadenylation and/or transcription termination in vivo.

The sequence discussed above contains the AATAAA sequence. To determine whether the AATAAA is required for 3′-end formation, we used site-directed mutagenesis to construct two additional plasmids. In pL107, the AATAAA sequence was changed to AATTCA, and in pL108, the AATAAA was deleted. In pL107, the 2-nt change in the sequence should be sufficient to inactivate the polyadenylation signal if the mechanism of 3′-end formation in yeasts is comparable to the reaction in mammalian cells. In support of this assumption, Wickens and Stephenson (49) have shown that both single mutations, AATACA and AATTAA, drastically reduce cleavage and polyadenylation of precursor RNA microinjected into *Xenopus* oocytes to 1.6 and 0.5% of normal levels, respectively. The data shown in Fig. 2B demonstrate that the double mutation and the deletion mutation did not abolish 3′-end formation activity, since only low levels of β-galactosidase activity were detected with both pL107 and pL108. This suggests that the AATAAA sequence is not absolutely required for 3′-end formation in *S. cerevisiae*.

To determine whether the position of the 3′ end was affected by these mutations, RNA was prepared from strains harboring plasmids with the mutations in AATAAA and compared with strains containing wild-type plasmids. The results from RNase protection analysis are shown in Fig. 4A. In this experiment, total RNA was isolated from the strains harboring pL101, pL102, pL107, or pL108. In pL101, pL102, and pL107, the RNA was hybridized to the antisense probe depicted in Fig. 3B, while in pL108, the RNA was hybridized to a probe derived from a similar antisense RNA which contained a deletion for AATAAA. The results demon-

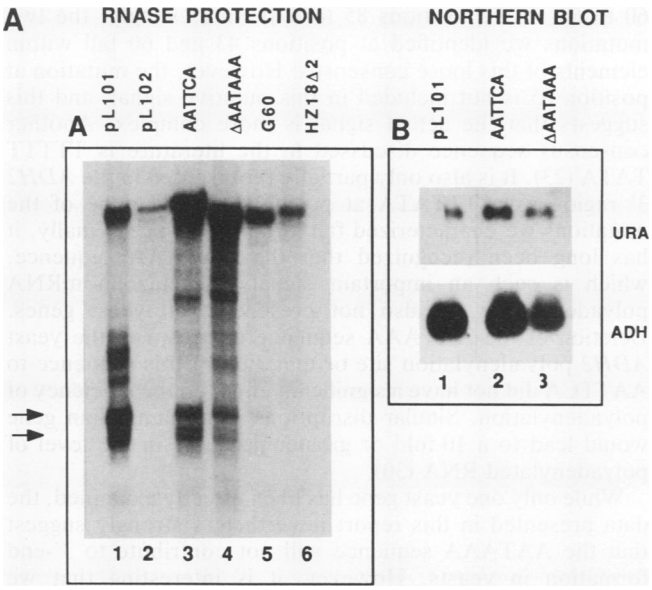

FIG. 4. Effect of AATAAA mutations on *ADH2* 3'-end formation. (A) Total RNA was prepared from strains harboring pL101 (lane 1), pL102 (lane 2), pL107, the plasmid containing AATTCA (lane 3), pL108, the plasmid containing a deletion of AATAAA (lane 4), pL401, the plasmid containing a point mutation which affects 3'-end formation (lane 5), and pHZ18Δ2, the parental plasmid lacking any *ADH2* sequence (lane 6). Each RNA was hybridized to an antisense *ADH2* probe and treated with RNase T2 before analysis on denaturing gels. Arrows point to major cleavage sites. (B) RNA was run on a 1.5% agarose-formaldehyde gel, and Northern blots were hybridized to either a *URA3* probe or an *ADH2* probe. Autoradiography was for 16 h (A) or 6 h (B).

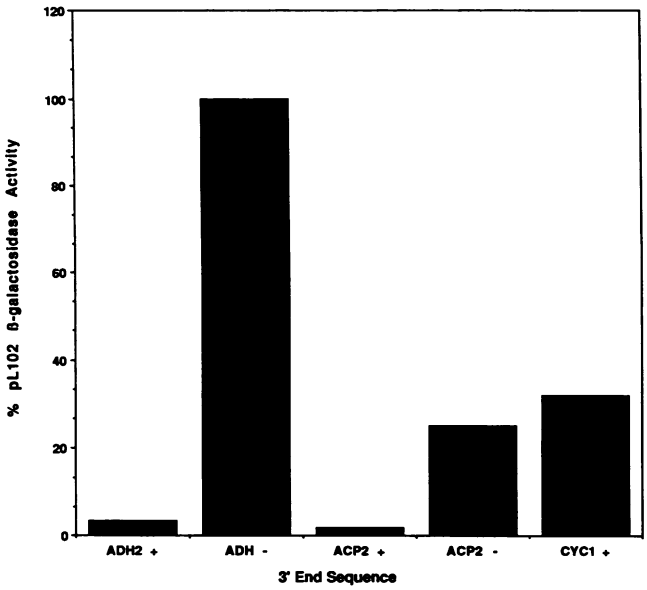

FIG. 5. Comparison of the efficiency of 3'-end sites by quantitative analysis of β-galactosidase activity. The relative amount of β-galactosidase activity was determined by analysis of solubilized cell extracts with *o*-nitrophenyl-β-D-galactopyranoside as the enzyme substrate. The data were compiled by averaging the results of approximately 10 independent experiments for each strain. +, Normal orientation; −, reverse orientation.

strated that the site of cleavage is unaltered by the mutations. However, the amount of RNA detected from the AATAAA deletion plasmid pL108 was reduced compared with that detected from the wild-type plasmid pL101 (Fig. 4A, lanes 1 and 4).

To establish the relative amounts of RNA derived from the various fusion gene plasmids, we analyzed RNA by Northern blots which were hybridized to either an *ADH2* probe or a probe complementary to the mRNA from the *URA3* gene, which serves as an internal control for the amount of RNA loaded onto the gel (Fig. 4B). In this and other experiments, we detected a slight reduction in the amount of RNA derived from the strain containing pL108 compared with the wild-type strain. More sensitive densitometry scanning of the gels demonstrated that this approximately twofold reduction in mRNA levels is consistent with the increase in the level of β-galactosidase observed for pL108 (Fig. 2B). Since the effect on 3'-end formation is rather subtle, it is most likely that the AATAAA sequence is not serving as a signal for cleavage and polyadenylation in yeasts as it is in metazoans.

We also compared the termination-polyadenylation activity of other 3'-end sequences inserted into the same site in the rp51 intron. The well-characterized *CYC1* 3' end (7, 40, 51) also directed 3'-end formation in this fusion gene, but much less efficiently than *ADH2* (Fig. 5). In this case, there was 35% readthrough of transcription, resulting in intermediate levels of β-galactosidase activity. This is the same relative amount of activity measured by Ruohola et al. (37). In their study, the *CYC1* 3'-end sequence was inserted into the actin gene intron, which was fused in frame to the *HIS4* gene, and readthrough of transcription was measured by quantitation of the resulting RNAs. We also tested sequences from the *ACP2* gene (20). Like the *ADH2* sequence, this sequence also functioned in an orientation-specific manner in that there was at least 10-fold-more β-galactosidase activity in the opposite orientation. However, there may be a weak signal in the negative orientation for this gene, since β-galactosidase levels were reduced compared with those generated by pHZ18Δ2 or pL102. The presence of overlapping polyadenylation signals may be a common feature of convergently transcribed yeast genes (29, 42, 46, 47, 51). It is also clear from these results that there is a wide variation in the efficiency at which specific sequences direct 3'-end formation.

The results described in the preceding paragraphs demonstrated that the AATAAA sequence is not required for 3'-end formation at the yeast *ADH2* site. To determine which other sequences in the *ADH2* 3' end are important for termination-polyadenylation, we mutagenized the plasmid pL101 in vitro, using hydroxylamine hydrochloride. After mutagenesis, the DNA was transformed into yeast cells and grown on indicator plates containing 5-bromo-4-chloro-3-indolyl-β-galactoside (X-gal). If a mutation results in inactivation of 3'-end formation, the resulting colony will appear blue on these plates. Mutations which detrimentally affect splicing, transcription, or translation of the fusion transcript will result in white colonies on indicator plates and will therefore be undetected in this screen. The sequencing of mutagenized DNA which inactivated the *ADH2* 3'-end signal(s) revealed a region 80 nt upstream from the cleavage site that is important for proper 3'-end formation. As depicted in Fig. 1 and 2A, pL401 contains a single change of T to G at position 60 (relative to the beginning of the *ADH2* insert). This mutation greatly increased the production of β-galactosidase (Fig. 2B) and caused a corresponding decrease in

the amount of RNA terminating at the *ADH2* site (Fig. 2C, lane 3). When the *ADH2* region from pL401 was excised from the plasmid and reinserted into pHZ18Δ2 Sma, the same results were obtained. Thus, the inactivation of the 3′-end-formation signal was due to this single change and was not due to another change elsewhere on the plasmid. We isolated three additional mutations, which are outlined in Fig. 1. The mutations in these plasmids all caused increased β-galactosidase activity when compared with the wild-type sequence (Fig. 2B). RNA analysis of these mutations demonstrated that there is a reduction of *ADH2* signal consistent with the increase in β-galactosidase activity (Fig. 2C). It is important to note that the mutations are clustered within 20 nt of each other. Interestingly, although they were induced by hydroxylamine, the observed changes include transversions, transitions, and deletions. Figure 2A shows the minimal *ADH2* sequence required to direct 3′-end formation and the mutations which affect this process. The pL402 mutation is a double mutation, resulting in the deletion of the T at position 60 and a C at position 65. The pL403 mutation is an G-to-A transition at position 45, and the pL404 mutation is a transversion at position 65 from C to G. Not all the mutations restored full β-galactosidase levels. For example, pL403, which contains the mutation furthest away from the others, resulted in the least severe phenotype. In this strain, β-galactosidase levels were reduced relative to those resulting from the other mutations, and the *ADH2* RNA levels were increased (Fig. 2C, lane 5). It is formally possible that these mutations increase the efficiency of splicing rather than increase the level of readthrough transcription. This possibility seems unlikely since the mutations are not located in regions known to be required for splicing. Nevertheless, we are in the process of testing this possibility by examining the levels of readthrough transcription directed by these *cis* mutations in a nonintron context.

DISCUSSION

Defining the sequences required for RNA-processing reactions has been a key factor in elucidating the mechanisms governing these reactions. It is well documented that the sequence AAUAAA and a U (or GU)-rich element downstream of this site are crucial to 3′-end formation of most eukaryotic mRNAs (30, 36). It is postulated that cleavage factors, a specificity factor, and the poly(A) polymerase all depend on these sequences for specific interaction with the precursor RNA (10, 11, 18, 31, 44). The purification of these factors, as well as the cloning of the genes coding for these factors, is an area undergoing intense investigation. We are expanding our understanding of how 3′-end formation occurs by examining the reaction in the yeast *S. cerevisiae*, a system amenable to genetic analysis.

An early step in defining the mechanism of 3′-end formation in *S. cerevisiae* is the identification of sequences required for the reaction. Although consensus sequences for polyadenylation signals in yeasts have been proposed (23, 51), these sequences are not uniformly present in yeast genes. For example, the sequence TAG. . .TA(T)GT. . . TTT may serve as an important signal for termination of transcription at some yeast sites (43, 51). This is based largely on a comparison of a limited number of yeast gene sequences and has not been directly tested by mutational analysis. An exact match to this sequence is not found in the *ADH2* 3′ end. However, there is a permutation of the sequence present in the region we defined as important for 3′-end formation (TAG [positions 43 to 45] TATG [positions 60 to 64] TTT [positions 85 to 87]). Interestingly, the two mutations we identified at positions 43 and 60 fall within elements of this loose consensus. However, the mutation at position 65 is not included in this putative signal, and this suggests that the actual signal is more complex. Another consensus sequence discussed in the literature is TTTTT TATA (23). It is also only partially represented in the *ADH2* 3′ region as a TTTATA at position 39, and none of the mutations we characterized fall in this sequence. Finally, it has long been recognized that the AATAAA sequence, which is such an important signal in metazoan mRNA polyadenylation, is also not present in all yeast genes. Deletion of the AATAAA sequence upstream of the yeast *ADH2* polyadenylation site or mutation of this sequence to AATTCA did not have a significant effect on the efficiency of polyadenylation. Similar disruptions in a mammalian gene would lead to a 10-fold or greater decrease in the level of polyadenylated RNA (30).

While only one yeast gene has been directly examined, the data presented in this report nevertheless strongly suggest that the AATAAA sequence will not contribute to 3′-end formation in yeasts. However, it is interesting that we detected a small increase in the amount of readthrough transcription seen in a gene which is missing this sequence. This result suggests that the deletion of the AATAAA sequence from the 3′ end of the *ADH2* gene does have a small effect on the efficiency but not the accuracy of 3′-end formation. Although it is unlikely that the AATAAA sequence plays the same role in polyadenylation of yeast mRNA as it does in metazoans, the data do not rule out the possibility that AATAAA is an important signal but that functionally redundant signals exist. Deletion of sequences critical for one pathway may activate a compensatory mechanism(s) for 3′-end formation.

It is somewhat surprising that AATAAA, or another consensus sequence, is not a central player in yeast 3′-end formation. Other basic reactions involved in mRNA synthesis, such as transcription (19) and pre-mRNA splicing (50), share very similar mechanisms between yeasts and higher eukaryotes. While the mutational analysis of the *ADH2* site revealed striking differences in the polyadenylation signals utilized by yeast and mammalian cells, it also highlighted some similarities. In both systems, there appears to be a requirement for a sequence upstream of the poly(A) addition site as well as one downstream. In the yeast *ADH2* gene, an upstream signal is located approximately 80 nt away from the cleavage site, compared with 10 to 30 nt for the mammalian signal. A similar situation exists for the yeast Ty-D15 element (47). In light of this, there is recent evidence that sequences upstream of the AATAAA are important for the efficient polyadenylation of some virus mRNAs (9, 16, 39) as well as some plant RNAs (24, 32).

The efficiency and precision of mRNA polyadenylation in yeasts resembles that of mammalian cells and logically implies the involvement of signal sequences on the RNA precursor. However, given the results presented here and the observations of others, there may not be a simple recognition sequence for 3′-end formation. Alternative hypotheses include highly degenerate *cis*-acting sequences, multiple sequence configurations which can lead to polyadenylation, or structural signals. A polyadenylation signal based on secondary structure may explain why it has not been possible to find a conserved consensus sequence for yeast polyadenylation. The *ADH2* sequence analyzed in this study can be folded into a stem-loop structure (shown in Fig. 2D). It is interesting that the sequences uncovered by the

hydroxylamine mutagenesis are clustered in a stem region. In addition, three of the mutations affect a bulged base in the stem structure. Site-directed mutagenesis experiments are currently under way to test the relevance of this stem-loop structure in the RNA. There is precedence for stem-loop structures as important signals for 3'-end formation in both eukaryotes and prokaryotes. In *E. coli*, the formation of a stem-loop structure in the RNA directs rho-independent termination. In this case, the stem-loop sequence is not conserved, but there is some sequence requirement in that the stem is usually followed by a stretch of U residues (35). Histone 3'-end formation in mammalian cells is another example of a well-defined system in which the RNA secondary structure is necessary for 3'-end formation (4). In this case, the signal specifying RNA processing is not the AAUAAA sequence but a stem-loop structure found in the 3' untranslated region of the RNA.

It is interesting that one of the mutations revealed by the hydroxylamine treatment does not result in a completely nonfunctional poly(A) site. If a factor which normally interacts at this site is weakly operative on the mutated site, increased expression of this putative factor(s) may allow suppression of the mutant phenotype. It is also possible that a compensatory mutation in a *trans*-acting factor could restore efficient 3'-end formation. A similar assumption was made in a study on yeast pre-mRNA splicing, and this approach was successful in the isolation and cloning of a *trans*-acting suppressor of a branchpoint mutation (6, 14). Thus, a search for suppressors of the mutations affecting polyadenylation may also lead to isolation of *trans*-acting factors required for 3'-end formation.

One of the fundamental conclusions of the data presented here is that the activity directed by the *ADH2* sequence is dominant over pre-mRNA splicing and that this causes the decrease in β-galactosidase activity. One possible explanation is that the *ADH2* sequence in the rp51 intron interferes somehow with recognition of splicing signals. Since the highly conserved yeast splicing signals, the GUAUGU at the 5' splice site and the UACUAAC at the branch site (50), are not found in the *ADH2* sequence, it is unlikely that this stretch of RNA is providing alternative signals which would lead to nonfunctional product RNAs. On longer exposures of Northern blots, such as in Fig. 2C and 4B, we did not observe longer species containing intron sequence with any of our constructs, as might be expected if splicing was inhibited. When yeast mRNA splicing is perturbed by *cis*- or *trans*-acting mutations which affect this processing reaction, unprocessed precursor and splicing intermediates accumulate in vivo and can be detected by such an analysis (14, 50). Instead, we saw only the short transcript correctly terminated at the *ADH2* site. Other studies also indicate that small insertions upstream of, but not adjacent to, the UACUAAC site of yeast introns have little effect on splicing (21, 27, 37).

If the *ADH2* site is a true processing site, it appears that the polyadenylation reaction preempts pre-mRNA splicing in yeasts. This is in direct contrast to the situation in higher eukaryotes. In this case, a poly(A) site placed within an intron is not recognized (2, 5, 28). In yeasts, the polyadenylation reaction may occur more rapidly than splicing, or formation of a 3'-end-processing complex may preclude assembly of the splicing apparatus. An alternative to this possibility is that the *ADH2* site is causing termination of transcription. If this is so, it is easy to reconcile the apparent discrepancies in the two systems, since the yeast fusion transcript would not include the 3' splice site.

We have not yet isolated any *cis*-acting mutations that allow us to separate the two major reactions contributing to the formation of mRNA 3' ends, termination of transcription and cleavage and polyadenylation. In this light, it is interesting that the *ADH2* and the *ACP2* mRNAs used in this study are very poor substrates for cleavage and polyadenylation in vitro (26a). This is somewhat unexpected given the efficiency at which these sites direct 3'-end formation in vivo. In contrast, the *CYC1* sequence is an efficient substrate for the cleavage and polyadenylation reaction in yeast extracts (7, 8) but is approximately threefold-less effective than the *ADH2* site for 3'-end formation in vivo. It will be interesting to establish whether the differences between the in vivo results and the in vitro activity reflect different mechanisms for 3'-end formation.

ACKNOWLEDGMENTS

We thank Dean Dawson, Fariba Houman, Marco Kessler, Claudio Pikielny, Anne Skvorak, and Andrew Wright for many helpful discussions and useful comments on the manuscript. Special thanks to Michele Flatters for constructing pL108, to Steve Schneider for excellent technical assistance, and to Andy Ellington for help with the RNA secondary structure analysis.

This work was supported by research fellowship award GM 12818-01A1 to L.E.H. and Public Health Service grant GM 41752-02 to C.L.M. from the National Institutes of Health and by grants from the Whittaker Health Science Fund, the Charlton Fund, and the Massachusetts Division of the American Cancer Society.

REFERENCES

1. **Abe, A., Y. Hiraoka, and T. Fukasawa.** 1990. Signal sequences for generation of mRNA 3' end in *Saccharomyces cerevisiae* GAL7 gene. EMBO J. **9**:3691–3697.
2. **Adami, G., and J. R. Nevins.** 1988. Splice site selection dominates over poly(A) site choice in RNA production from complex adenovirus transcription units. EMBO J. **7**:2107–2116.
3. **Ausubel, F. M., R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
4. **Birchmeier, C., W. Folk, and M. Birnstiel.** 1983. The terminal RNA stem-loop structure and 80bp of spacer DNA are required for the formation of 3' termini of sea urchin H2A mRNA. Cell **35**:433–440.
5. **Brady, H. A., and W. S. M. Wold.** 1988. Competition between splicing and polyadenylation reactions determines which adenovirus region E3 mRNAs are synthesized. Mol. Cell. Biol. **8**:3291–3297.
6. **Burgess, S., J. R. Couto, and C. Guthrie.** 1990. A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. Cell **60**:705–717.
7. **Butler, J. S., and T. Platt.** 1988. RNA processing generates the mature 3' ends of yeast CYC1 mRNA in vitro. Science **242**: 1270–1274.
8. **Butler, J. S., 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.
9. **Carswell, S., and J. C. Alwine.** 1989. Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. Mol. Cell. Biol. **9**:4248–4258.
10. **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.
11. **Christofori, G., and W. Keller.** 1989. Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre-mRNA in vitro. Mol. Cell. Biol. **9**:193–203.
12. **Chu, C. T., D. S. Parris, R. A. Dixon, F. E. Farber, and P. A. Scaffer.** 1979. Hydroxylamine mutagenesis of HSV DNA and DNA fragments: introduction of mutations into selected regions of the viral genome. Virology **98**:168–181.
13. **Connelly, S., and J. L. Manley.** 1988. A functional mRNA

polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. **2**:440–452.
14. **Couto, J. R., J. Tamm, R. Parker, and C. Guthrie.** 1987. A trans-acting suppressor restores splicing of a yeast intron with a branch point mutation. Genes Dev. **1**:445–455.
15. **Devereux, J. R., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the Vax. Nucleic Acids Res. **12**:387–395.
16. **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.
17. **Falck-Pedersen, E., J. Logan, T. Shenk, and J. E. Darnell.** 1985. Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse β-globin terminator element. Cell **40**:897–905.
18. **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.
19. **Guarante, L.** 1988. UASs and enhancers: common mechanism of transcriptional activation in yeast and mammals. Cell **52**:303–305.
20. **Haggren, W., and D. Kolodrubetz.** 1988. The *Saccharomyces cerevisiae ACP2* gene encodes an essential HMG1-like protein. Mol. Cell. Biol. **8**:1282–1289.
21. **Halfter, H., and D. Gallwitz.** 1988. Impairment of yeast pre-mRNA splicing by potential secondary structure-forming sequences near the conserved branchpoint sequence. Nucleic Acids Res. **16**:10413–10423.
22. **Hart, R. P., M. A. McDevitt, H. Ali, and J. R. Nevins.** 1985. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and the simian virus 40 polyadenylation sites. Mol. Cell. Biol. **5**:2975–2983.
23. **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.
24. **Hunt, A. G., and M. H. MacDonald.** 1989. Deletion analysis of the polyadenylation signal of a pea ribulose-1,5-bisphosphate carboxylase small-subunit gene. Plant Mol. Biol. **13**:125–138.
24a. **Hyman, L., and C. Moore.** Unpublished observations.
25. **Inouye, S., and M. Inouye.** 1987. Oligonucleotide-directed site-specific mutagenesis using double-stranded plasmid DNA, p. 181–206. *In* S. Narang (ed.), DNA and RNA synthesis. Academic Press, Inc., New York.
26. **Joshi, C. P.** 1987. Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. Nucleic Acids Res. **15**:9627–9640.
26a. **Kessler, M., and C. Moore.** Unpublished data.
27. **Klinz, F.-J., and D. Gallwitz.** 1985. Size and position of intervening sequences are critical for the splicing efficiency of pre-mRNA in the yeast *Saccharomyces cerevisiae*. Nucleic Acids Res. **13**:3791–3804.
28. **Levitt, N., D. Briggs, A. Gil, and N. Proudfoot.** 1989. Definition of an efficient synthetic poly(A) site. Genes Dev. **3**:1019–1025.
29. **Liu, Y., and C. L. Dieckmann.** 1989. Overproduction of yeast viruslike particles by strains deficient in a mitochondrial nuclease. Mol. Cell. Biol. **9**:3323–3331.
30. **Manley, J. L.** 1988. Polyadenylation of mRNA precursors. Biochim. Biophys. Acta **950**:1–12.
31. **McDevitt, M. A., G. M. Gilmartin, W. H. Reeves, and J. R. Nevins.** 1988. Multiple factors are required for poly(A) addition to a mRNA 3′ end. Genes Dev. **2**:588–597.
32. **Mogen, B., M. H. MacDonald, R. Graybosch, and A. G. Hunt.** 1991. Upstream sequences other than AAUAAA are required for efficient messenger RNA 3′-end formation in plants. Plant Cell **2**:1261–1272.
33. **Osborne, B. I., and L. Guarante.** 1989. Mutational analysis of a yeast transcriptional terminator. Proc. Natl. Acad. Sci. USA **86**:4097–4101.
34. **Pikielny, C. W., J. L. Teem, and M. Rosbash.** 1983. Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implication for U1 RNA and metazoan mRNA splicing. Cell **34**:395–403.
35. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
36. **Proudfoot, N. J.** 1989. How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. **14**:105–110.
37. **Ruohola, H., S. M. Baker, R. Parker, and T. Platt.** 1988. Orientation-dependent function of a short CYC1 DNA fragment in directing mRNA 3′ end formation in yeast. Proc. Natl. Acad. Sci. USA **85**:5041–5045.
38. **Russell, D. W., M. Smith, V. M. Williamson, and E. T. Young.** 1983. Nucleotide sequence of the yeast alcohol dehydrogenase II gene. J. Biol. Chem. **258**:2674–2682.
39. **Russnak, R., and D. Ganem.** 1990. Sequences 5′ to the polyadenylation signal mediate differential poly(A) site use in hepatitis B virus. Genes Dev. **4**:764–776.
40. **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.
41. **Sherman, F., G. R. Fink, and J. B. Hicks.** 1986. Laboratory course manual for methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
42. **Snyder, M., R. J. Sapolsky, and R. W. Davis.** 1988. Transcription interferes with elements important for chromosome maintenance in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:2184–2194.
43. **Sutton, A., and J. R. Broach.** 1985. Signals for transcription initiation and termination in the *Saccharomyces cerevisiae* plasmid 2 μm circle. Mol. Cell. Biol. **5**:2770–2780.
44. **Takagaki, Y., L. C. Ryner, and J. L. Manley.** 1988. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA processing and polyadenylation. Cell **52**:731–742.
45. **Teem, J. L., and M. Rosbash.** 1983. Expression of a β-galactosidase gene containing the ribosomal protein 51 intron is sensitive to the rna2 mutation of yeast. Proc. Natl. Acad. Sci. USA **80**:4403–4408.
46. **Yarger, J. G., G. Armilei, and M. C. Gorman.** 1986. Transcription terminatorlike element within a *Saccharomyces cerevisiae* promoter region. Mol. Cell. Biol. **6**:1095–1011.
47. **Yu, K., and R. 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.
48. **Whitelaw, E., and N. Proudfoot.** 1986. Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3′ end processing in the human α-globin gene. EMBO J. **5**:2915–2922.
49. **Wickens, M., and P. Stephenson.** 1984. Role of the conserved AAUAAA sequence: four point mutants prevent messenger RNA 3′ end formation. Science **226**:1045–1051.
50. **Woolford, J. L.** 1989. Nuclear pre-mRNA splicing in yeast. Yeast **5**:439–457.
51. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28**:563–573.
52. **Zucker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.