4676–4687 *Nucleic Acids Research, 1998, Vol. 26, No. 20*

Regulation of poly(A) site choice of several yeast mRNAs

Kimberly A. Sparks[1] and Carol L. Dieckmann[1,2,*]

[1]Department of Biochemistry and [2]Department Molecular and Cellular Biology, Life Sciences South Room 454, University of Arizona, Tucson, AZ 85721, USA

Received June 16, 1998; Revised and Accepted August 31, 1998

ABSTRACT

Several yeast genes produce multiple transcripts with different 3′-ends. Of these, four genes are known to produce truncated transcripts that end within the coding sequence of longer transcripts: *CBP1*, *AEP2*/*ATP13*, *RNA14* and *SIR1*. It has been shown that the level of the truncated *CBP1* transcript increases during the switch to respiratory growth while that of the full-length transcript decreases. To determine whether this phenomenon is unique to *CBP1*, northern analysis was used to determine whether the levels of other truncated transcripts are regulated similarly by carbon source. The levels of the shortest transcripts of *AEP2*/*ATP13* and *RNA14* increased during respiration while the shortest *SIR1* transcript remained constant. However, two longer *SIR1* transcripts were regulated reciprocally by carbon source. Mapping the 3′-ends of each transcript by sequencing partial cDNA clones revealed multiple 3′-ends for each transcript. Examination of the sequences surrounding the 3′-ends of the induced transcripts failed to identify a consensus sequence but did reveal weak putative 3′-end formation signals in all of the transcripts. Similarly, no consensus sequence was found when the sequences surrounding the 3′-ends of the longest transcripts were compared, but again weak putative 3′-end formation signals were identified. These data are suggestive of carbon source regulation of alternative poly(A) site choice in yeast.

INTRODUCTION

In many types of animal cells (mammalian, avian, amphibian, insect, fish and trypanosome) and several animal viruses, alternative poly(A) site choice has been described as a method for regulating gene expression. A recent review described 126 genes in which alternative 3′-end formation, sometimes in combination with alternative splicing, resulted in formation of multiple mRNAs from the same transcription unit (1) Of these 126 genes, there are at least 33 examples of differential regulation of poly(A) site choice at different developmental stages or in different tissues (1). Some well-studied examples include cyclin D1, CD40, eIF-4E, dihydrofolate reductase, herpes simplex virus type 1 (HSV-1) UL24 and iron regulatory protein 2 (for a review see 1). The abundance of examples of alternative polyadenylation illustrates the importance of poly(A) site choice in the regulation of gene expression in animal cells.

In yeast cells, only one example of regulated alternative poly(A) site choice has been described, that of *CBP1* (2). Two types of transcripts are produced from the *CBP1* template; these transcripts share a common 5′-end but differ at the 3′-ends (2). The longer transcript encodes a 66 kDa mitochondrial protein that is required for accumulation of mature mitochondrial cytochrome *b* mRNA and thus is essential for respiration (3–7). Western analysis failed to detect a protein translated from the shorter transcript (8). Disruption of the long transcript by insertion of *LEU2* near the end of the *CBP1* coding sequence renders yeast cells respiratory-deficient (2), therefore, the short transcript alone is insufficient to support respiration. Since the protein encoded by the long transcript is required for respiration, it was surprising that the steady-state level of the long transcript decreased upon induction of respiration while that of the short transcript increased (2). Since total *CBP1* transcript levels remain constant during induction of respiration (9), it is unlikely that transcription induction is involved in the regulation of *CBP1* transcript levels.

Several lines of evidence were suggestive that the levels of the two types of *CBP1* transcripts are reciprocally regulated by carbon source by alternative 3′-end formation. First, insertion of a 146 bp fragment of *CBP1* surrounding the 3′-ends of the short transcript into a reporter gene (*URA3*) resulted in production of a new, shorter transcript; the levels of the two *URA3* transcripts were reciprocally regulated by carbon source in a manner similar to that of the two types of *CBP1* transcripts (9). This result is suggestive that the short transcript 3′-end formation element is necessary and sufficient for carbon source-dependent regulation of short transcript production. In another study, measurement of the decay rates of the two types of *CBP1* transcripts in fermenting or respiring cells revealed that carbon source-dependent differences in mRNA degradation rates could not explain reciprocal regulation of the two types of *CBP1* transcripts (10). In a third study, an *in vitro* transcriptional pausing assay failed to detect pausing of RNA polymerase II in the short transcript 3′-end formation element (R.Weilbaecher and C.M.Kane, personal communication). Similarly, deletion of the gene encoding the transcription elongation factor TFIIS, which helps polymerase II read-through

*To whom correspondence should be addressed at: Department of Biochemistry, Life Sciences South Room 454, University of Arizona, 1007 East Lowell, Tucson, AZ 85721, USA. Tel: +1 520 621 3569; Fax: +1 520 621 9903; Email: dieckman@u.arizona.edu

pause sites (for a review see 11), had no effect on short or long *CBP1* transcript production *in vivo* (10). Thus, transcriptional pausing is not responsible for the regulation of *CBP1* transcript levels. A final study found that mutation of the short transcript 3′-end formation signal eliminated short transcript production and also abolished the carbon source-dependent decrease in the level of the long transcript (8). Collectively, these data are supportive of the hypothesis that carbon source-dependent alternative poly(A) site choice regulates the levels of the two types of *CBP1* transcripts.

We wondered whether carbon source-regulated alternative poly(A) site choice was a more general phenomenon during induction of respiration. It is known that several other yeast nuclear genes, in addition to *CBP1*, produce multiple transcripts with common 5′-ends but different 3′-ends. These genes include *GAL1* (12), *URA3* (13), *ALG7* (14), *MCM1* (15),(16), *MOD5* (17), *SUA7* (B.C.Hoopes, personal communication), *RNA14* (18), *SIR1* (19) and *AEP2*/*ATP13* (20). Of these, *RNA14*, *SIR1* and *AEP2*/*ATP13* produce truncated transcripts that have 3′-ends within the coding sequence of longer transcripts, similar to the short *CBP1* transcript. Like *CBP1*, *AEP2*/*ATP13* is necessary for respiration; Aep2/Atp13 is required for production of the mitochondrially encoded *ATP9* mRNA at a post-transcriptional step (20). Thus, it would not be surprising if the levels of the two types of *AEP2*/*ATP13* transcripts are regulated by carbon source in a manner similar to that of the two types of *CBP1* transcripts. Rna14, however, is a component of the general 3′-end formation machinery (21), while Sir1 represses transcription from the silent mating type loci (19). Therefore, there is no reason to expect that *RNA14* or *SIR1* transcript levels will be regulated by carbon source. Also, the 3′-ends of the long and short transcripts of *AEP2*/*ATP13*, *RNA14* and *SIR1* have not been accurately mapped, so it has not been possible to compare the sequences surrounding the 3′-ends of the short or long transcripts for conserved motifs that might be responsible for differential 3′-end formation. In this study, northern analysis was used to determine whether the levels of the multiple transcripts produced by *AEP2*/*ATP13*, *RNA14* and *SIR1* are regulated during the switch to respiratory growth in the same way as the levels of the two types of *CBP1* transcripts. The 3′-end of each transcript made from all four genes was mapped by sequencing partial cDNA clones. Two types of sequence comparison were used to search for a consensus sequence common to the induced transcripts but absent from the longest transcripts, which were not induced, that might be a binding site for a carbon source-specific regulatory protein. In addition, possible 3′-end formation signals for each type of truncated transcript were identified.

MATERIALS AND METHODS

Strains and media

The *Saccharomyces cerevisiae* and *Escherichia coli* strains and plasmids used in this study are listed in Table 1. To repress respiration, yeast cells were grown overnight without agitation to a low density in minimal salts medium (0.3% yeast extract, 1.2% NH_4SO_4, 0.7% $MgSO_4$, 1% KH_2PO_4, 0.5% NaCl, 0.4% $CaCl_2$) (22) containing 10% glucose. To induce respiration, the cells were grown with vigorous agitation in minimal salts medium containing 5% glycerol. *Escherichia coli* strains were grown in LB medium (23); ampicillin was added to a final concentration of 100 μg/ml where required. Solid media contained 2% agar.

Table 1. Yeast and *E.coli* strains and plasmids

Strain or plasmid	Genotype	Source or reference
Yeast strains		
N5-26/LA1	MAT**a** *cbp1-1 ade1 leu2-3,112*	2
T31	N5-26/LA1 transformed with pG60/T31	2
S150-2B	MAT**a** *ura3-52 his3Δ leu2-3,112 trp1-289*	8
E.coli strains		
RR1	F^- *hsdS20*(r^B m^B) *ara-14 proA2 lacY1 galK2 rpsL20*(Sm^r) *xyl-5 mtl-1 supE44* l^- *recA1 endA1 gyrA96 thi hsdR17*($r_K^-m_K^+$) *relA1 lac* [F′ *proAB* $lacI^q$ZDM15 Tn*10*(*tet*)]	49
Yeast plasmids		
pG60/T31	2μ plasmid containing a 2.4 kb *Sau3*A fragment of chromosome X, including *CBP1*	50
pG95/ST3–2	Yep352 containing a 4 kb *Xba*I fragment which includes *AEP2*/*ATP13*	S.Ackerman
E.coli plasmids		
Bluescript KS+	T3 and T7 RNA polymerase promoters flank the multiple cloning site in the KS orientation, with the f1 origin in the + orientation	Stratagene
pBS+	T3 and T7 RNA polymerase promoters flank the multiple cloning site, with the f1 origin in the + orientation	Stratagene
pUC19	*E.coli* vector containing a multiple cloning site	51
pBS1-2	pBS+ with the 700 bp *Bam*HI–*Hin*dIII fragment of *CBP1* inserted	2
pKS/ACT1$^+$	Bluescript KS+ containing the ~600 bp *Cla*I fragment of *ACT1*	6
pLP75	Bluescript KS+ containing a 2.9 kb *Sma*I–*Hin*dIII fragment which includes *SIR1*	E.Stone
pLM4	pUC19 containing a *Hin*dIII fragment which includes the coding sequence of *RNA14*	E.Mandart

RNA isolation and northern blot analysis

Total RNA was isolated by the glass beads method from aliquots of yeast strain T31 frozen at various time points after induction of respiration as described previously (2). RNA concentration was determined by measurement of A_{260}. Poly(A)$^+$ RNA was enriched on oligo(dT)–cellulose columns as described previously (2) and ~12 µg of each sample were separated by non-denaturing electrophoresis on a vertical agarose gel. The gel was soaked in 50 mM sodium hydroxide, stained with ethidium bromide in TB (0.083 M Tris base, 0.089 M boric acid) and destained in TB. Northern blot analysis was performed exactly as described previously (8).

Preparation of probes

The *CBP1* cRNA probe CBP1-2 (2), the *ACT1* cRNA probe pKS/ACT1$^+$ (6) and the *SIR1* cRNA probe pG95/ST3-2 were radiolabeled with [α-^{32}P]UTP (ICN, Costa Mesa, CA) by *in vitro* transcription using T3 RNA polymerase (Boehringer Mannheim, Indianopolis, IN) according to the manufacturer's instructions. A 520 bp *Eco*RI–*Pst*I fragment of pLM4 (*RNA14*) and a 1100 bp *Bgl*II–*Sca*I fragment of pLP75 (*AEP2/ATP13*) were gel purified and uniformly labeled with [α-^{32}P]dATP (ICN) using the Random Primed DNA Labeling Kit (Boehringer Mannheim) according to the manufacturer's instructions. Quantitation of RNA levels was performed using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). The transcript levels shown in Figure 1 were derived by subtracting the background level from the signal for each transcript (including *ACT1*) and then dividing the corrected value by the level of *ACT1* mRNA, which is constant during induction of respiration (2), for each lane of the gel. Transcript levels at the 0 h time point were then set to 1.00 for each probe and the remaining values were adjusted accordingly.

Mapping of mRNA 3′-ends

All enzymes and restriction endonucleases were purchased from Boehringer Mannheim. Total RNA from aliquots of S150-2B frozen at various time points after induction of respiration was prepared by the hot phenol method as described previously (8); poly(A)$^+$ RNA was then enriched by oligo(dT)–cellulose batch preps as described previously (8). Partial cDNA clones containing the 3′-ends of each transcript made from *CBP1*, *AEP2/ATP13*, *SIR1* and *RNA14* were isolated from poly(A)$^+$ RNA as described in the RACE protocol (24), with a PCR annealing temperature of 54°C, using AMV reverse transcriptase and *Taq* DNA polymerase. All oligonucleotides were synthesized by National Biosciences (Plymouth, MN) or Genosys (Woodlands, TX); the sequences of the oligonucleotides are listed in Table 2. The first strand cDNA synthesis (3′) primer contained restriction sites for *Sal*I, *Xba*I and *Eco*RI and a run of 17 T residues. The corresponding 3′ PCR adapter primer was identical except that it lacked the $(T)_{17}$ extension. The unincorporated nucleotides and mineral oil were removed from the PCR products with the QIAquick Spin PCR Purification Kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. The purified PCR products were digested with *Eco*RI (in the 3′ primer) and the following restriction enzymes: 1.2 kb *CBP1*, *Hin*dIII (starting at +793); 2.2 kb *CBP1*, *Bam*HI (in the 5′ primer); 0.6 kb *AEP2/ATP13*, *Xho*I (starting at +321); 1.85 and 2.2 kb *SIR1*, *Kpn*I (in the 5′ primer); 2.3 and 2.4 kb *SIR1*, *Bam*HI (in the 5′ primer); 1.1 and 1.5 kb *RNA14*, *Bam*HI (in the 5′ primer); 2.2 kb *RNA14*, *Xba*I (starting at +1623). The purified PCR products for the 2.1 kb *AEP2/ATP13* transcript were digested with *Xba*I (in the 3′ primer) and *Eco*RI (starting at +1553). The fragments were gel purified and ligated to appropriately digested Bluescript KS+ using standard techniques (23).

Escherichia coli was transformed with the ligation reaction products using the previously described Hanahan transformation procedure (25). Plasmid inserts were sequenced from miniprep DNA using T7 and T3 primers and the Sequenase v.2.0 DNA Sequencing Kit (Amersham, Arlington Heights, IL), except in cases where Sequenase did not proceed through the run of 17 T residues. Thus, the following primers were also used for sequencing: *RNA14* 1.1 kb transcript clones, *RNA14*+879 primer; *RNA14* 1.5 kb transcript clones, *RNA14*+1280 primer; *SIR1* 1.85 kb transcript clones, *SIR1*+1627 primer; *AEP2/ATP13* 2.1 kb transcript clones, *ATP13*+1777 primer. The following are the number of clones sequenced for each transcript: *CBP1* 1.2 kb, 10; *CBP1* 2.2 kb, 6; *AEP2/ATP13* 0.6 kb, 10; *AEP2/ATP13* 2.1 kb, 8; *RNA14* 1.1 kb, 10; *RNA14* 1.5 kb, 12; *RNA14* 2.2 kb, 9; *SIR1* 1.85 kb, 13; *SIR1* 2.2 kb, 15; *SIR1* 2.3 kb, 3; *SIR1* 2.4 kb, 8.

RESULTS

The levels of several different types of truncated transcripts are regulated by carbon source

The levels of the two types of *CBP1* transcripts are regulated reciprocally by carbon source. When yeast cells are induced to respire by growth on a non-fermentable carbon source such as glycerol, the steady-state level of the full-length transcript decreases while that of the truncated transcript increases (2). We hypothesized that truncated transcripts produced by other genes might be regulated similarly. *AEP2/ATP13* (20), *RNA14* (18) and *SIR1* (19) all produce short transcripts with 3′-ends in the coding sequence of full-length transcripts (Fig. 1). To determine whether the levels of these other short transcripts are regulated by carbon source in the same way as the short *CBP1* transcript, poly(A)-enriched RNA was isolated from yeast strain T31 at 4 h intervals after induction of respiration and analyzed by northern blot via successive hybridization with probes specific for *CBP1*, *AEP2/ATP13*, *RNA14* and *SIR1* (Fig. 2). The quantitation of these four experiments is shown in Figure 1.

As observed previously (2,8,9), the steady-state level of the 1.2 kb *CBP1* transcript increased upon induction of respiration (~2-fold) while the level of the 2.2 kb *CBP1* transcript decreased reciprocally (Figs 1 and 2A). Like *CBP1*, *AEP2/ATP13* produces two types of transcripts: full-length 2.1 kb and truncated 0.6 kb transcripts. Also similar to *CBP1*, the steady-state level of the 0.6 kb short transcript of *AEP2/ATP13* increased upon induction of respiration (~7-fold), while the steady-state level of the 2.1 kb long transcript decreased reciprocally (Figs 1 and 2B). Regulation of the two types of *AEP2/ATP13* transcript by carbon source was not surprising since this gene is necessary for respiration.

Unlike *CBP1* and *AEP2/ATP13*, three types of transcript are produced from the *RNA14* template: full-length 2.2 kb and truncated 1.5 and 1.1 kb transcripts. Like the *CBP1* and *AEP2/ATP13* short transcripts, the steady-state level of the 1.1 kb short transcript of *RNA14* also increased upon induction of respiration (~14-fold), while the steady-state levels of the 1.5 kb truncated and the 2.2 kb long transcripts decreased (Figs 1 and 2C). This result is exciting because *RNA14* has no known respiratory function. It is also interesting that only the shortest (1.1 kb) transcript increased in abundance during respiration while the level of the other truncated (1.5 kb) transcript decreased, like that of the full-length (2.2 kb) transcript.

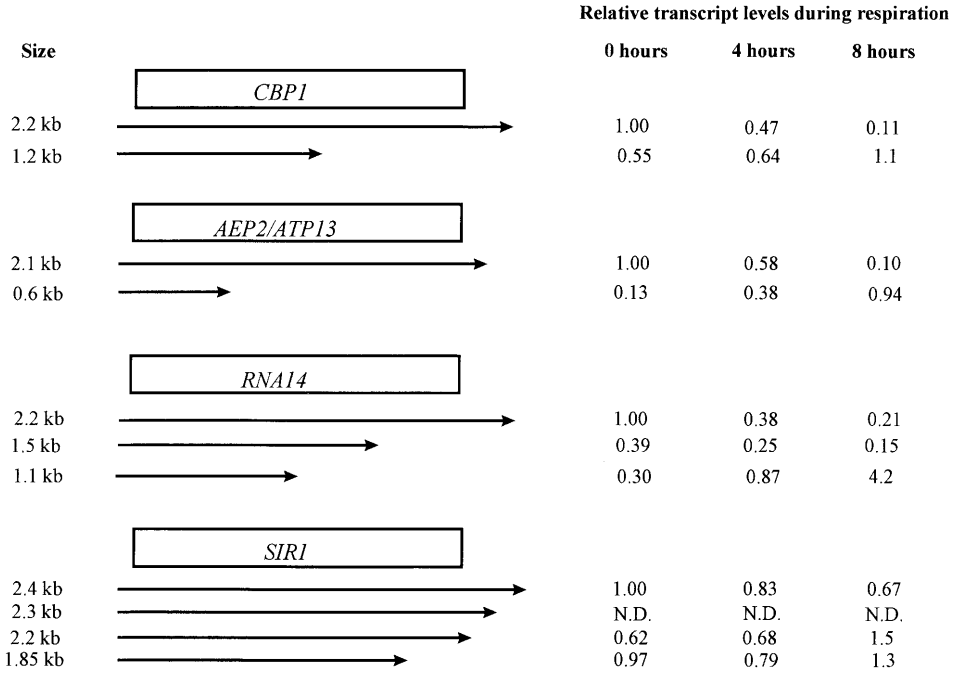


Figure 1. Schematic representation of the transcripts produced by *CBP1* (2), *AEP2/ATP13* (20), *RNA14* (18) and *SIR1* (19). The transcripts produced by each gene share common 5′-ends but differ at the 3′-ends. In all four cases, the longest transcript encodes the protein that carries out the function of the gene while the shorter transcript(s) has 3′-ends within the coding sequence of the long transcript. The only exceptions are the 2.2 and 2.3 kb *SIR1* transcripts, which end in the 3′-UTR. The boxes represent the coding sequence of each gene, while the arrows represent the transcripts. The 2.2 and 2.3 kb transcripts of *SIR1* have not been previously reported. The relative transcript levels during respiration represent a quantitation of the bands observed in Figure 2. Transcript levels were adjusted for background and normalized to *ACT1* mRNA levels. The long transcript level at the 0 h time point for each gene was set to 1.00 and the remaining values were adjusted accordingly. Units are arbitrary. Times represent hours after induction of respiration. N.D., not determined.

Table 2. Oligonucleotides used in this study

Name	Sequence	Description
3′ cDNA	GCGTCGACTCTAGAGAATTC(T)$_{17}$	First strand cDNA synthesis primer
PCR adapter	GCGTCGACTCTAGAGAATTC	PCR adapter primer (3′)
CBP1+700	AAGGCCCTCGGCTTAAGTAATG	*CBP1* +700 to +720
CBP1+1900	CGGGATCCAACAAGGAGATCATACAAGATA	*CBP1* +1900 to +1921 with a *Bam*HI site added to the 5′-end
ATP13+280	GGTATACCTATAGATGTTCACA	*ATP13* +280 to + 301
ATP13+1551	GGGAATTCATTGATAGTTATGG	*ATP13* +1551 to +1572
ATP13+1777	CTCCATATTCTTATTCTATAG	*ATP13* +1777 to +1797
SIR1+1386	GGGGTACCGGAAGACCTTTTGATTACTCT	*SIR1* +1386 to +1406 with a *Kpn*I site added to the 5′-end
SIR1+1627	TATCAAGATCCTATCCCG	*SIR1* +1627 to +1644
SIR1+2025	GGGGTACCCTGTAACGATTAAATAGTTGGTA	*SIR1* +2025 to +2047 with a *Kpn*I site added to the 5′-end
SIR1+2075	GGGGATCCTAAGCGGGTGATGTCTCATTT	*SIR1* +2075 to +2095 with a *Bam*HI site added to the 5′-end
RNA14+793	CGGGATCCCCACAACCAGGTACCTCAG	*RNA14* +793 to +811 with a *Bam*HI site added to the 5′-end
RNA14+879	GCTTAGTGAAGATATGCT	*RNA14* +879 to +896
RNA14+1127	CGGGATCCCTCAAACTTTACTATCGCAG	*RNA14* +1127 to +1146 with a *Bam*HI site added to the 5′-end
RNA14+1280	CGGGATCCCAGGACTATCCGCAGCAC	*RNA14* +1280 to +1297 with a *Bam*HI site added to the 5′-end
T3	ATTAACCCTCACTAAAG	T3 promoter primer
T7	AATACGACTCACTATAG	T7 promoter primer

All oligonucleotides are listed 5′→3′.

In the case of *SIR1*, two types of transcripts had been detected previously: full-length 2.4 kb and truncated 1.7 kb transcripts (here measured as 1.85 kb; below). Unlike the short transcripts produced by *CBP1* and *AEP2/ATP13* and the shortest *RNA14* transcript, the steady-state level of the 1.85 kb short transcript of *SIR1* remained constant upon induction of respiration (Figs 1 and 2D). Interestingly, a new type of *SIR1* long transcript (~2.2 kb) was detected using a *SIR1* antisense riboprobe; the 2.2 kb transcript is predicted to end beyond the stop codon of *SIR1*. As seen for *CBP1*, *AEP2/ATP13* and *RNA14*, the level of the longer, 2.4 kb transcript decreased during induction of respiration while the level of the shorter, 2.2 kb transcript increased (>2-fold). Collectively, these data are suggestive that carbon source-regulated alternative 3′-end formation might be a more general phenomenon in yeast than was originally assumed. The results of the northern analyses are also suggestive of two classes of truncated transcripts, those that are induced during respiratory growth (the 1.2 kb *CBP1* transcript, the 0.6 kb *AEP2/ATP13* transcript, the 1.1 kb *RNA14* transcript and the 2.2 kb *SIR1* transcript) and those that are not induced (the 1.5 kb *RNA14* transcript and the 1.85 kb *SIR1* transcript).

Mapping the 3′-ends of the short and long transcripts of *CBP1, AEP2/ATP13, RNA14* and *SIR1* reveals that each transcript has multiple 3′-ends

We hypothesized that the transcripts that are induced might have a *cis*-element required for increased 3′-end formation during induction of respiration. In order to deduce such a *cis*-element from sequence alignments, it was first necessary to map the 3′-ends of the short and long transcripts of *AEP2/ATP13*, *RNA14* and *SIR1*. Mapping the 3′-ends of the two types of long *SIR1* transcripts would also confirm that these transcripts have different 3′-ends. The 3′-ends of the short and long transcripts of *CBP1* were mapped as a positive control, since these ends had been mapped previously (9). The 3′-ends of the transcripts were mapped by sequencing partial cDNA clones of each transcript obtained by the RACE method of RT-PCR (24). First-strand cDNA synthesis was performed using a 3′ primer complementary to the poly(A) tail (3′ cDNA) which added a 'tag' to the products (Table 2). The products were amplified by PCR using a 3′ primer that contained the tag sequence (PCR adapter) and gene-specific 5′ sense primers from sequence that was estimated to be 200–400 bp upstream of the 3′-ends (Table 2). Figure 3 summarizes the locations of the 3′-ends of the 11 classes of transcripts made by the four genes studied. All of the transcripts have multiple 3′-ends within a short region. The consensus sequence for yeast poly(A) addition sites is $Py(A)_n$ (where Py represents a pyrimidine) (26–28). However, these 11 classes of transcripts were often polyadenylated at non-consensus sites.

Interestingly, mapping the 3′-ends of the *SIR1* transcripts revealed three classes of long transcripts differing by ~100 bp, in addition to the short transcript. The position of the 3′-end of the short *SIR1* transcript is suggestive that this transcript is closer to 1.8–1.9 than 1.7 kb. The sizes of the three longest transcripts are estimated to be 2.2, 2.3 and 2.4 kb based on the positions of the 3′-ends. The three longest transcripts appeared as two barely resolved bands on a northern blot (Fig. 2D). We assume that the highest band on the northern blot is of a size to be comprised of the 2.4 kb transcript and the smaller band is of a size to be comprised of the 2.2 kb transcript. The 2.3 kb transcript was not observed running between the other two. This may be due to the lower abundance of this size class of transcript or perhaps it was not resolved on this gel. Therefore, potential regulation of the 2.3 kb *SIR1* transcript could not be determined and this transcript was excluded from the sequence comparisons below.

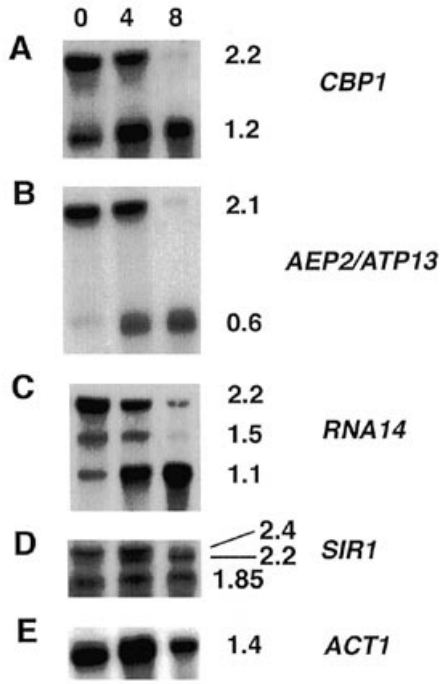


Figure 2. Northern analysis reveals that the production of several truncated transcripts is regulated by carbon source. Yeast cells were induced to respire and poly(A)$^+$ RNA isolated at 4 h intervals after initiation of induction was separated by non-denaturing electrophoresis, transferred to Nytran and hybridized with successive probes, stripping the blot between probes. The numbers above the lanes represent hours after induction of respiration while the numbers to the right of the lanes represent transcript sizes (in kb). This figure was made using Adobe Photoshop v.3.0. (**A**) *CBP1*; (**B**) *AEP2/ATP13*; (**C**) *RNA14*; (**D**) *SIR1*; (**E**) *ACT1*, which was used to normalize for differences in loading.

Collectively, mapping the 3′-ends of 11 types of transcripts produced by the four genes revealed multiple 3′-ends for each class of transcript; similar results were seen previously for genes that produce only one transcript (28–32). In addition, the 3′-ends of each type of transcript studied here were distributed over regions spanning from 4 to >200 nt. Mapping the 3′-ends of the long and short transcripts produced by *CBP1*, *AEP2/ATP13*, *RNA14* and *SIR1* allowed sequence comparisons to try to elucidate the mechanism by which truncated transcript levels are regulated by carbon source.

Sequence comparisons fail to reveal a consensus sequence in the induced transcripts

We propose two models to explain the regulation of alternative 3′-end formation by carbon source (Fig. 4). The specific factor model states that a specific regulatory protein binds to the RNA surrounding the cleavage site of each induced transcript to increase (if the protein is an activator) or decrease (if the protein is a repressor) production of that transcript in response to carbon source. This model predicts the existence of a consensus binding

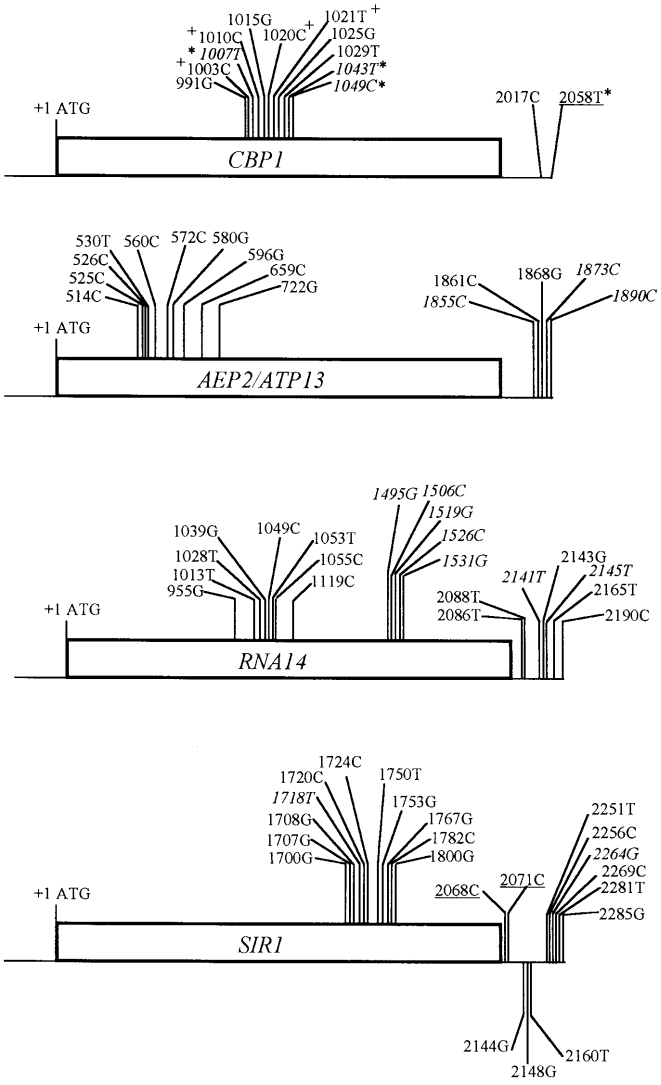


Figure 3. Location of the 3′-ends of the long and short transcripts produced by *CBP1, AEP2/ATP13, RNA14* and *SIR1*. The 3′-ends of each transcript were mapped by sequencing partial cDNA clones obtained by the RACE method (24). Numbers are relative to the ATG of each gene at +1. The asterisks represent previously mapped *CBP1* 3′-ends that were also isolated in this study, while the crosses represent *CBP1* 3′-ends that were only isolated in the previous study (9). Italicized text represents 3′-ends isolated two or three times, while underlined text represents 3′-ends isolated six or more times.

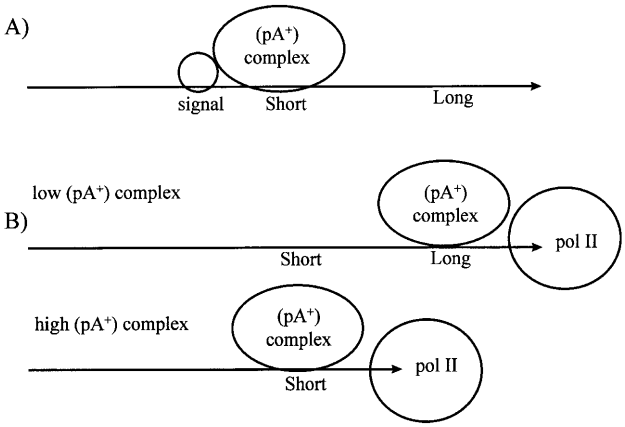


Figure 4. Two models of the regulation of alternative poly(A) site choice of several yeast transcription units. (**A**) Specific factor model. A specific regulatory protein (small circle) binds to a conserved element (signal) in the RNA to nucleate binding of the cleavage/polyadenylation complex [oval labeled (pA^+) complex] to the upstream poly(A) site (short). (**B**) General machinery model. At low levels of cleavage/polyadenylation activity [(low pA^+) complex], the complex would bind to the higher affinity, downstream signal (long). A general increase in 3′-end formation activity [(high pA^+) complex] would then allow the complex to bind the lower affinity, upstream sites (short) also and cleave the transcript before the downstream signal is transcribed, since cleavage occurs co-transcriptionally. Note that both the cleavage/polyadenylation complex and RNA polymerase II (pol II) are multisubunit enzymes and are depicted as a single protein only to simplify the drawing.

site (either a sequence motif or a structure) for the regulatory protein in the induced transcripts. Alternatively, the general machinery model states that an increase in the amount or activity of one of the general 3′-end formation factors upon induction of respiration results in increased use of weak upstream 3′-end formation signals such as those found in the induced transcripts studied here. This model does not predict a consensus protein binding site but instead predicts that the 3′-end formation signals of the induced transcripts will not be optimal (i.e. they will contain mismatches to the proposed 3′-end formation consensus sequences).

In order to test the specific factor model, we searched for possible binding sites for a putative carbon source-dependent regulatory protein in the induced transcripts. The sequences surrounding the mapped 3′-ends of the four types of induced transcripts (the 0.6 kb *AEP2/ATP13* transcript, the 1.2 kb *CBP1* transcript, the 1.1 kb *RNA14* transcript and the 2.2 kb *SIR1* transcript) were compared using the Pileup program (Wisconsin Package v.9.0; Genetics Computer Group, Madison, WI). Although sequence identities were observed at many positions throughout the alignment, the sequence comparison failed to reveal an obvious consensus sequence (data not shown). After the sequences were aligned, the sequence of the MMH2 allele of *CBP1* (which contains 50 silent mutations in the region surrounding the 3′-ends of the short transcript and does not make the short transcript) (8) was added manually by aligning the MMH2 allele sequence with the corresponding sequence of the wild-type *CBP1* allele. While the sequence of the MMH2 allele was different from the weak, assigned consensus in many positions, the differences were spread throughout the sequence and did not help to elucidate a conserved element.

Since no consensus binding site for an activator of short transcript 3′-end formation was revealed by comparing the sequences upstream of the 3′-ends of the induced transcripts, we wondered if a sequence comparison of the four types of longest transcripts (called ‘terminal transcripts’) that decrease during the switch to respiration might help to elucidate a consensus binding site in the induced transcripts. In other words, we wanted to look for differences in the alignment of the terminal transcripts from the alignment of the induced transcripts to try to find a consensus that was not apparent from studying only the alignment of the induced transcripts. The Pileup program was used to align the sequence surrounding the 3′-ends of the 2.2 kb *CBP1* transcript, the 2.1 kb *AEP2/ATP13* transcript, the 2.4 kb *SIR1* transcript and the 2.2 kb *RNA14* transcript. Again, no obvious consensus was revealed; the weak assigned consensus differed from that of the

induced transcripts at many positions (data not shown). Therefore, alignment of the four types of terminal transcripts did not help to locate a possible protein binding site in the induced transcripts. The failure to detect a strongly conserved sequence in these two different sequence alignments is suggestive that there is not a specific sequence common to these four types of induced transcripts that is responsible for binding a carbon source-specific regulatory protein.

Since the Pileup program failed to identify a consensus sequence by searching for sequence conservation among the four types of induced transcripts, we next looked for a conserved sequence element located at a conserved distance upstream of the 3′-end of each induced transcript. Only sequences upstream of the 3′-ends were examined because no essential downstream element for yeast 3′-end formation has been described (33). The Lineup program of Wisconsin Package v.9.0 was used to align the 60 nt sequences upstream of each poly(A) addition site of the four types of induced transcripts. The site of poly(A) addition is located immediately to the left of the vertical line for each sequence (Fig. 5A). The Lineup program also assigned a consensus at each position representing the most common nucleotide present at that position. Examination of the computer-assigned consensus sequence revealed a very A+T-rich sequence, with only three C residues and no G residues out of 54 assigned positions. However, examination of the sequence of each transcript reveals that each nucleotide was allowed at each position (supported by the low percentage conservation at each position), suggestive that no absolute sequence is required at any position located a certain distance from the poly(A) addition site.

Further examination of the sequence upstream of each poly(A) addition site revealed the presence of potential 3′-end formation signals for each transcript. These signals are known to be functionally weak because none of the induced transcripts are efficiently formed and thus longer transcripts are made from each gene. Yeast 3′-end formation signals are thought to consist of three distinct elements: the efficiency element (consensus sequence UAUAUA, UAUGUA or UUUUUAUA), the positioning element (consensus sequence AAGAA, AAAAAA or AAUAAA) and the poly(A) addition site itself [consensus sequence Py(A)$_n$] (34). The putative 3′-end formation signals of the four types of induced transcripts contain mismatches to the proposed consensus sequences for efficiency and positioning elements, which could lead to the observed heterogeneity of 3′-ends for each type of truncated transcript and the functional weakness of the signal (34). Also, only 55% of the cleavage sites were of the consensus Py(A)$_n$ type. Perhaps the suboptimal spacing of the positioning and efficiency elements with respect to each other and to the poly(A) addition site also contributes to the inefficiency of the signals.

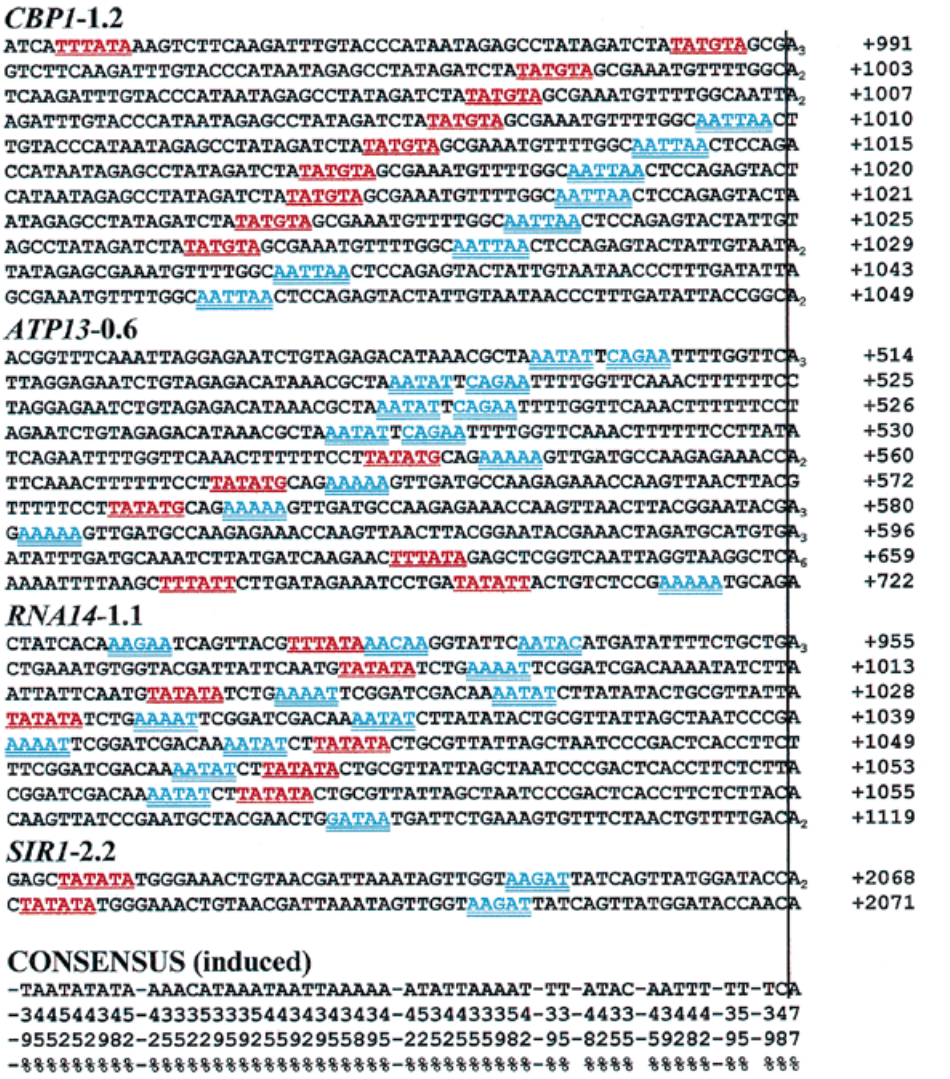
To determine whether the putative 3′-end formation elements for the 1.2 kb *CBP1* transcript were altered in making the MMH2 allele, a similar Lineup was performed using the MMH2 sequence corresponding to the sequences used for the wild-type *CBP1* transcripts in Figure 5A. The Lineup of the MMH2 sequences revealed mutations in all of the proposed 3′-end formation signals for the 1.2 kb *CBP1* transcript (Fig. 5B). Since the MMH2 allele of *CBP1* does not allow production of short transcripts (8), the presence of mutations in the putative 3′-end formation signals is suggestive that these non-optimal signals may be responsible for 1.2 kb *CBP1* transcript formation.

To look for 3′-end formation signals in the four types of longest transcripts that decrease during respiration, the Lineup program was again used to align the 60 nt sequences upstream of each poly(A) addition site of the 2.2 kb *CBP1* transcript, the 2.1 kb *AEP2/ATP13* transcript, the 2.2 kb *RNA14* transcript and the 2.4 kb *SIR1* transcript. The computer-assigned consensus sequence for the regions immediately upstream of the four types of terminal transcripts was found to be very A+T-rich (no C or G residues out of 55 assigned positions; Fig. 5C), similar to the consensus sequence for the four types of induced transcripts (compare with Fig. 5A). Another similarity to the induced transcripts was the presence of potential 3′-end formation signals in the sequence immediately upstream of the poly(A) addition sites of most of the transcripts; most of the potential signals contain mismatches to the proposed consensus signals. The presence of non-consensus signals was surprising since no longer transcripts are formed from each transcription unit, suggestive that these signals must be completely efficient. Additionally, 80% of the cleavage sites are of the Py(A)$_n$ type, suggestive that these transcripts have functionally stronger 3′-end formation signals than the induced truncated transcripts. Comparison of the consensus sequences for the induced and terminal transcripts reveals many similarities (Fig. 5D). However, one major difference is the high A-richness of the 28 nt upstream of the cleavage site in the terminal consensus. This region is much less A-rich in the induced consensus, suggestive that this region is important in making a strong 3′-end formation signal. Perhaps this A run represents a strong positioning element for the terminal transcripts, resulting in efficient use of the most distal 3′-end formation site in each transcription unit under fermenting (non-inducing) conditions.

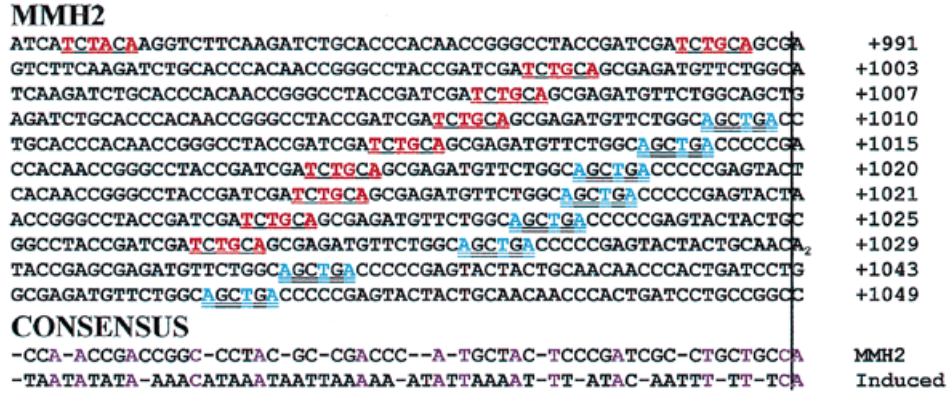
A similar Lineup was performed for the constitutively expressed 1.85 kb *SIR1* transcript (Fig. 5E). The consensus sequence was then compared with the consensus sequences for the induced and terminal transcripts. The consensus sequence for the 1.85 kb *SIR1* transcript is more A-rich than that of the induced transcript but less A-rich than that of the terminal transcript. The A-rich region of the *SIR1* 1.85 kb consensus is also shifted upstream with respect to the A-rich region in the terminal consensus. Perhaps the lower A content and the greater distance from the 3′-ends of the 1.85 kb *SIR1* transcript results in the observed high incidence of cleavage after G residues, which was not seen in either the induced or the terminal transcripts. However, perhaps the signal is strong enough to allow efficient 3′-end formation of the 1.85 kb *SIR1* transcript such that the level of this transcript remains constant upon induction of respiration (see Figs 1 and 2D).

In studying the Lineups for all the transcripts analyzed here, we were generally able to identify efficiency elements with one or fewer mismatches to the proposed consensus in all of the transcripts studied. However, we noticed that it was difficult to identify sequences in most of the transcripts that match the proposed consensus sequences for positioning elements of yeast 3′-end formation signals (34). The only transcripts with appropriately placed positioning elements that match the proposed consensus are the 2.2 kb *RNA14* transcript and the 1.85 kb *SIR1* transcript. It is especially striking that the terminal transcripts, which are known to have functionally strong 3′-end formation signals, do not have consensus-type efficiency and positioning elements. Therefore, we propose that, while the consensus efficiency and positioning elements are both necessary and sufficient for 3′-end

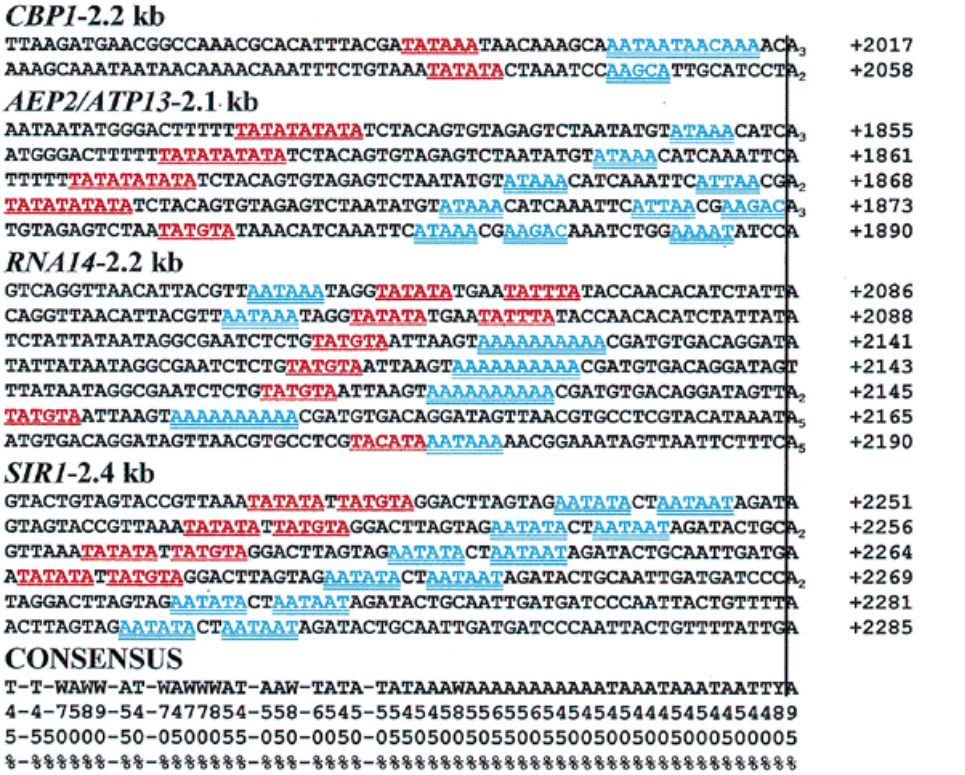
A


CBP1-1.2
+991
+1003
+1007
+1010
+1015
+1020
+1021
+1025
+1029
+1043
+1049
ATP13-0.6
+514
+525
+526
+530
+560
+572
+580
+596
+659
+722
RNA14-1.1
+955
+1013
+1028
+1039
+1049
+1053
+1055
+1119
SIR1-2.2
+2068
+2071
CONSENSUS (induced)
-TAATATATA-AAACATAAATAATTAAAAA-ATATTAAAAT-TT-ATAC-AATTT-TT-TCA
-344544345-4333533354434343434-4534433354-33-4433-43444-35-347
-955252982-2552295925592955895-2252555982-95-8255-59282-95-987
-%%%%%%%%%-%%%%%%%%%%%%%%%%%%%%%-%%%%%%%%%%-%% %%%% %%%%%-%% %%%

B


MMH2
+991
+1003
+1007
+1010
+1015
+1020
+1021
+1025
+1029
+1043
+1049
CONSENSUS
-CCA-ACCGACCGGC-CCTAC-GC-CGACCC--A-TGCTAC-TCCCGATCGC-CTGCTGCCA MMH2
-TAATATATA-AAACATAAATAATTAAAAA-ATATTAAAAT-TT-ATAC-AATTT-TT-TCA Induced

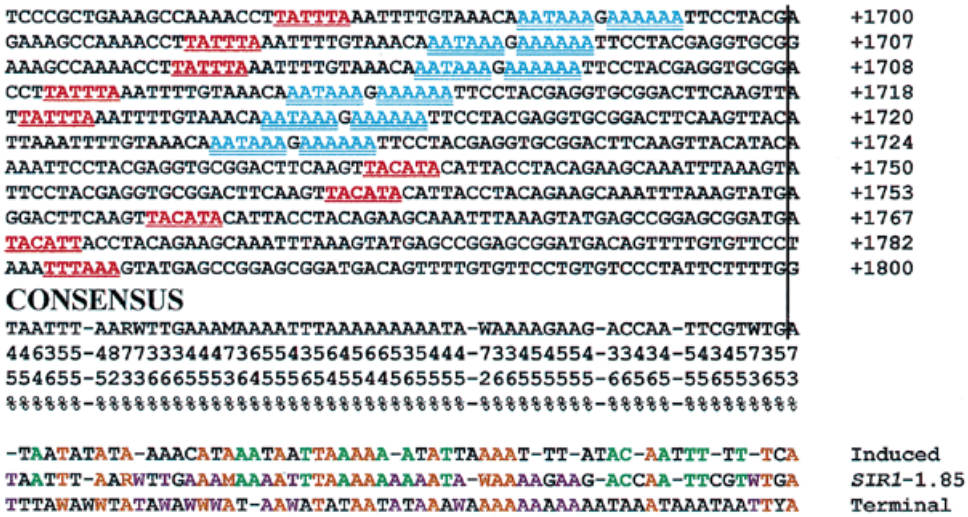
C


CBP1-2.2 kb
TTAAGATGAACGGCCAAACGCACATTTACGATATAAATAACAAAGCAAATAATAACAAAACA3 +2017
AAAGCAAATAATAACAAAACAAATTTCTGTAAATATATACTAAATCCAAGCATTGCATCCTA2 +2058
AEP2/ATP13-2.1 kb
AATAATATGGGACTTTTTTATATATATATCTACAGTGTAGAGTCTAATATGTATAAACATCA3 +1855
ATGGGACTTTTTTATATATATATCTACAGTGTAGAGTCTAATATGTATAAACATCAAATTCA +1861
TTTTTTATATATATATCTACAGTGTAGAGTCTAATATGTATAAACATCAAATTCATTAACGA2 +1868
TATATATATATCTACAGTGTAGAGTCTAATATGTATAAACATCAAATTCATTAACGAAGACA3 +1873
TGTAGAGTCTAATATGTATAAACATCAAATTCATAAACGAAGACAAAATCTGGAAAATATCCA +1890
RNA14-2.2 kb
GTCAGGTTAACATTACGTTAATAAATAGGTATATATGAATATTTATACCAACACATCTATTA +2086
CAGGTTAACATTACGTTAATAAATAGGTATATATGAATATTTATACCAACACATCTATTATA +2088
TCTATTATAATAGGCGAATCTCTGTATGTAATTAAGTAAAAAAAAAACGATGTGACAGGATA +2141
TATTATAATAGGCGAATCTCTGTATGTAATTAAGTAAAAAAAAACGATGTGACAGGATAGT +2143
TTATAATAGGCGAATCTCTGTATGTAATTAAGTAAAAAAAAACGATGTGACAGGATAGTTA2 +2145
TATGTAATTAAGTAAAAAAAAACGATGTGACAGGATAGTTAACGTGCCTCGTACATAAATA5 +2165
ATGTGACAGGATAGTTAACGTGCCTCGTACATAAATAAACGGAAATAGTTAATTCTTTCA5 +2190
SIR1-2.4 kb
GTACTGTAGTACCGTTAAATATATATTATGTAGGACTTAGTAGAATATACTAATAATAGATA +2251
GTAGTACCGTTAAATATATTATGTAGGACTTAGTAGAATATACTAATAATAGATACTGCA2 +2256
GTTAAATATATTATGTAGGACTTAGTAGAATATACTAATAATAGATACTGCAATTGATGA +2264
ATATATATTATGTAGGACTTAGTAGAATATACTAATAATAGATACTGCAATTGATGATCCCA2 +2269
TAGGACTTAGTAGAATATACTAATAATAGATACTGCAATTGATGATCCCAATTACTGTTTTA +2281
ACTTAGTAGAATATACTAATAATAGATACTGCAATTGATGATCCCAATTACTGTTTTATTGA +2285
CONSENSUS
T-T-WAWW-AT-WAWWWAT-AAW-TATA-TATAAAWAAAAAAAAAAAATAAATAAATAATTYA
4-4-7589-54-7477854-558-6545-554545855655654545454454454454489
5-550000-50-0500055-050-0050-055050050550055005005005000500005
%-%%%%%%-%%-%%%%%%%-%%%-%%%%-%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

D


-TAATATATA-AAACATAAATAATTAAAAA-ATATTAAAAT-TT-ATAC-AATTT-TT-TCA Induced
T-T-WAWW-AT-WAWWWAT-AAW-TATA-TATAAAWAAAAAAAAAAATAAATAAATAATTYA Terminal

E


SIR1-1.85
TCCCGCTGAAAGCCAAAACCTTATTTAAATTTTGTAAACAAATAAAGAAAAAATTCCTACGA +1700
GAAAGCCAAAACCTTATTTAAATTTTGTAAACAAATAAAGAAAAAATTCCTACGAGGTGCGG +1707
AAAGCCAAAACCTTATTTAAATTTTGTAAACAAATAAAGAAAAAATTCCTACGAGGTGCGGA +1708
CCTTATTTAAATTTTGTAAACAAATAAAGAAAAAATTCCTACGAGGTGCGGACTTCAAGTTA +1718
TTATTTAAATTTTGTAAACAAATAAAGAAAAAATTCCTACGAGGTGCGGACTTCAAGTTACA +1720
TTAAATTTTGTAAACAAATAAAGAAAAAATTCCTACGAGGTGCGGACTTCAAGTTACATACA +1724
AAATTCCTACGAGGTGCGGACTTCAAGTTACATACATTACCTACAGAAGCAAATTTAAAGTA +1750
TTCCTACGAGGTGCGGACTTCAAGTTACATACATTACCTACAGAAGCAAATTTAAAGTATGA +1753
GGACTTCAAGTTACATACATTACCTACAGAAGCAAATTTAAAGTATGAGCCGGAGCGGATGA +1767
TACATTACCTACAGAAGCAAATTTAAAGTATGAGCCGGAGCGGATGACAGTTTTGTGTTCCT +1782
AAATTTTAAAGTATGAGCCGGAGCGGATGACAGTTTTGTGTTCCTGTGTCCCTATTCTTTTGG +1800
CONSENSUS
TAATTT-AARWTTGAAAMAAAATTTAAAAAAAAATA-WAAAAGAAG-ACCAA-TTCGTWTGA
446355-48773334447365543564566535444-733454554-33434-543457357
554655-52336665553645556545544565555-266555555-66565-556553653
%%%%%%-%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%-%%%%%%%%%-%%%%%-%%%%%%%%%
-TAATATATA-AAACATAAATAATTAAAAA-ATATTAAAAT-TT-ATAC-AATTT-TT-TCA Induced
TAATTT-AARWTTGAAAMAAAATTTAAAAAAAAATA-WAAAAGAAG-ACCAA-TTCGTWTGA SIR1-1.85
TTTAWAWWTATAWAWWWAT-AAWATATAATATAAAWAAAAAAAAAAATAAATAAATAATTYA Terminal

formation of some yeast transcripts (26,27,32,35,36), other signals must also exist that function equally well.

DISCUSSION

The data presented here show that the steady-state levels of four different types of truncated transcripts (the 1.2 kb *CBP1* transcript, the 0.6 kb *AEP2*/*ATP13* transcript, the 1.1 kb *RNA14* transcript and the 2.2 kb *SIR1* transcript) increased during induction of respiration. It is not surprising that carbon source affects the levels of the short transcripts of *CBP1* and *AEP2*/*ATP13* since both of these genes have essential functions for respiration. However, we do not understand why the levels of the long, protein-encoding transcripts of *CBP1* and *AEP2*/*ATP13* decrease at a time when the proteins are needed (the switch to respiration). Also, *RNA14* and *SIR1* have no known respiratory functions and so there is no apparent reason for them to respond to changes in carbon source. The increase in short transcript levels for all four genes might be a transient response to the change in growth conditions. Indeed, switching yeast cells from rich glucose medium to a variety of nutritional and temperature conditions causes a switch in the levels of the two *SUA7* (yeast TFIIB) transcripts (B.C.Hoopes, personal communication), similar to that described here for *CBP1*, *AEP2*/*ATP13*, *RNA14* and *SIR1*.

Two of the genes studied here, *RNA14* and *SIR1*, produce more than two size classes of transcripts. Interestingly, the levels of the different classes of transcripts made by the two genes are regulated differently. In the case of *RNA14*, only the level of the shortest (1.1 kb) transcript increases upon induction of respiration at the expense of the two longer transcripts. In other words, when yeast cells are switched to respiratory growth, the pattern of use of the different 3′-end formation signals shifts from favoring the most promoter-distal signal to favoring the most promoter-proximal signal. The same type of shift is observed in the case of the two types of *CBP1* transcripts and the two types of *AEP2*/*ATP13* transcripts.

However, induction of respiration has very little effect on the level of the shortest (1.85 kb) *SIR1* transcript. Instead, the level of the 2.2 kb truncated transcript increases during respiration at the expense of the 2.4 kb full-length transcript and perhaps the 2.3 kb transcript as well. In other words, regulation of the longer transcripts is independent of the production of the shortest transcript. Perhaps the 1.85 kb *SIR1* short transcript lacks a *cis*-acting signal that is present in the 1.2 kb *CBP1* short transcript, the 1.1 kb *RNA14* short transcript, the 0.6 kb *AEP2*/*ATP13* short transcript and the 2.2 kb *SIR1* truncated transcript. However, no consensus sequence was found using two types of sequence comparison. Also, the Foldrna program of Wisconsin Package v.9.0 revealed potential stem–loops in the 60 nt sequence upstream of the 3′-end of each transcript (with no conservation of either sequence or position; data not shown); whether these stem–loops actually fold and are functionally important remains to be tested. In every case, the levels of truncated transcripts from four different genes increased during respiration at the expense of longer transcripts produced from the same gene. Perhaps truncated transcripts produced by other genes might also be induced by a switch to respiratory growth.

Taken together, our data are suggestive that the ‘specific factor’ model requiring a *cis* consensus site and a carbon source-regulated *trans* factor is not viable. We propose instead that increased short transcript levels are due to increased use of the short transcript 3′-end formation signals by the general 3′-end formation machinery. It has been shown that 3′-end cleavage and polyadenylation occur before transcription termination (37). Therefore, an increase in 3′-end cleavage would favor use of the weaker upstream 3′-end formation signals over the stronger downstream 3′-end formation signals because the upstream signals would be transiently more abundant than the downstream signals. Binding of the 3′-end formation complexes to the upstream polyadenylation site would allow cleavage at that site before the downstream site is transcribed. One inefficient 3′-end formation signal would be used to direct cleavage and polyadenylation at multiple positions for each type of truncated transcript.

If the general machinery model is true, we would predict that the use of any upstream 3′-end formation signal would increase during respiration due to increased 3′-end formation activity. However, it was observed that the level of the shortest type of *SIR1* transcript did not change during induction of respiration; interestingly, this transcript has putative 3′-end formation signals that more closely match the proposed consensus sequences than those of the other types of truncated transcripts. Perhaps this signal is efficient enough to be recognized when the activity of the 3′-end formation machinery is lower (during fermentation) such that an increase in activity would not significantly increase use of this site. However, this most upstream site must not be an optimal site because longer transcripts are also made from the *SIR1* template. Regulation of the levels of the three longer *SIR1* transcripts fits this model, since the most upstream of the three sites is used more often than the two downstream sites during respiration.

While few examples of regulated poly(A) site choice have been described in yeast (2; this study), many examples have been described in higher eukaryotes (reviewed in 1).

Figure 5. (Opposite and previous page). Comparison of the sequence upstream of the 3′-ends of the induced or repressed transcripts using the Lineup program of Wisconsin Package v.9.0. For each transcript, the 60 nt sequence immediately upstream of each poly(A) addition site was aligned by placing the poly(A) site immediately to the left of the vertical line, which represents the cleavage site. The numbers represent the nucleotide position of the poly(A) addition site relative to the ATG of each gene at +1. The consensus was assigned by the computer by determining the most frequent nucleotide at each position. A dash in the consensus sequence represents a position at which no consensus could be assigned. The numbers underneath the consensus represent the percentage of total sequences that have the consensus nucleotide at that position; numbers should be read vertically. (**A**) Lineup analysis of the four types of transcripts that are induced during respiration. Single underlined red sequences represent possible efficiency elements and double underlined blue sequences represent possible positioning elements of yeast 3′-end formation signals. (**B**) Lineup analysis of the sequence of the MMH2 allele of *CBP1* at the positions corresponding to the wild-type *CBP1* sequence in (A). Black nucleotides within the red or blue regions represent mutations in possible efficiency (red) or positioning (blue) elements for the wild-type 1.2 kb *CBP1* transcript. Purple nucleotides in the consensus represent positions where the MMH2 consensus matches the consensus for the induced transcript. (**C**) Lineup analysis of the four types of terminal transcripts. Symbols are as described in (A). (**D**) Comparison of the consensus sequences for the induced and the terminal transcripts. Differences between the two sequences are highlighted in orange. (**E**) Lineup analysis of the 1.85 kb *SIR1* transcript. The consensus for this transcript is compared with the consensuses for induced and terminal transcripts. Green nucleotides represent identity between the *SIR1* 1.85 kb consensus and the induced consensus, while purple nucleotides represent identity between the *SIR1* 1.85 kb consensus and the terminal consensus. Orange nucleotides represent identity between all three consensuses. Other symbols are as described in (A).

The general machinery model is supported by several known cases in mammalian cells. First, use of the weaker upstream 3′-end formation signal of the adenovirus major late transcript decreases late in infection. This decrease in short transcript formation is correlated with a decrease in activity of the 64 kDa subunit of the general 3′-end formation factor CstF (38), which binds to the downstream GU-rich element of the 3′-end formation signal in the nascent transcript (39,40). In another example supportive of the general machinery model, use of the weaker upstream Ig γ2b 3′-end formation signal increases during B cell development when the RNA-binding activity of CstF-64K increases (41). Overexpression of CstF-64K is sufficient to increase use of the weaker upstream Ig γ2b poly(A) site (42). While no similar examples have been described in yeast, it is reasonable to assume that the activity of the general yeast 3′-end formation machinery could be regulated by an important cellular process such as the switch to respiratory growth.

Based on the mammalian examples, there are several good candidates for 3′-end formation factors that might regulate truncated transcript formation upon induction of respiration. The most obvious candidates are the subunits of CF I, the complex that is thought to bind to the positioning element of yeast 3′-end formation signals (43). (The positioning element determines the site at which 3′-end formation occurs; 26,27.) CF I can be further separated into CF IA and CF IB (43). The known subunits of CF IA include Pcf11 (44), Clp1 (45), Pab1 (45), Rna14 and Rna15 (21). It was shown previously that temperature-sensitive alleles of *RNA14*, *RNA15* and *PAP1* [which encodes poly(A) polymerase] result in a shift from use of the most proximal 3′-end of actin at the permissive temperature to use of the most distal 3′-end, located ~180 bp downstream, at the restrictive temperature (46). This result is suggestive that loss of Rna14, Rna15 or Pap1 activity shifts the cleavage site downstream; perhaps increased activity of these proteins would shift the cleavage site upstream. The Rna15 protein shows sequence homology to the 64 kDa subunit of the mammalian 3′-end formation factor CstF (47). Rna15 contains RNP1 and RNP2 consensus RNA-binding domains (18) and can bind to *GAL7* RNA lacking an efficiency element (43); thus, Rna15 is thought to be the protein that recognizes and binds to the positioning element in nascent transcripts. Therefore, perhaps the Rna15 protein level or RNA-binding activity increases during the shift to respiration to allow increased recognition of the weak upstream sites.

In an attempt to assimilate the decrease in long *RNA14* transcripts upon induction of respiration into our general machinery model, we propose that the immediate decrease in growth rate when the cells are switched to respiratory conditions would quickly result in an increase in CF I specific activity on a per cell basis if the synthesis rates of the subunits remained the same for some time after the shift. An increase in CF I activity would result in increased recognition of weak but promoter-proximal 3′-end formation signals such as those found in the shortest transcripts of *CBP1*, *RNA14* and *AEP2/ATP13* and the 2.2 kb truncated transcript of *SIR1*. This increased activity would then result in a decrease in the level of the Rna14-encoding 2.2 kb transcript, which would in turn decrease the concentration of Rna14 protein. A decrease in the level of Rna14 would then result in decreased CF I activity, completing a negative feedback loop. An alternative hypothesis is that CF I activity increases due to an increase in the level or activity of other CF I proteins, despite lower amounts of Rna14.

An alternative version of the general machinery model is that one of the general 3′-end formation factors acts as a repressor of weak 3′-end formation signals. The activity of this repressor would decrease upon induction of respiration, thus allowing increased use of weak signals. One candidate for such a repressor is Hrp1/Nab4, which has been shown to be the sole component of CF IB (48). Hrp1/Nab4 was shown to repress use of weak, cryptic, upstream poly(A) sites in the *CYC1* transcript; interestingly, the nuclear concentration of Hrp1/Nab4 decreases during induction of respiration (A.M.Krecic and L.Minvielle-Sebastia, personal communication). The effects of Hrp1/Nab4 on the induced transcripts studied here remain to be determined. Similarly, many aspects of the general machinery model remain to be tested.

ACKNOWLEDGEMENTS

The authors wish to thank the following people: Elisabeth Mandart and François Lacroute for the gift of plasmid pLM4; Elisa Stone and Lorraine Pillus for the gift of plasmid pLP75; Sharon Ackerman for the gift of plasmid pG95/ST3-2; Barbara Hoopes, Annette M. Krecic and Lionel Minvielle-Sebastia for sharing unpublished results; Steve Mayer for preparing the poly(A)$^+$ RNA used in the northern blot analysis; Cathy Kingdon Worley for help with Figure 5; Doug Roberts and Laurie Marnell for critical reading of the manuscript. This study was supported by National Institutes of Health grant GM34893 to C.L.D.

REFERENCES

1 Edwalds-Gilbert,G., Veraldi,K.L. and Milcarek,C. (1997) *Nucleic Acids Res.*, **25**, 2547–2561.
2 Mayer,S.A. and Dieckmann,C.L. (1989) *Mol. Cell. Biol.*, **9**, 4161–4169.
3 Chen,W. and Dieckmann,C.L. (1994) *J. Biol. Chem.*, **269**, 16574–16578.
4 Dieckmann,C.L., Koerner,T.J. and Tzagoloff,A. (1984) *J. Biol. Chem.*, **259**, 4722–4731.
5 Mittelmeier,T.M. and Dieckmann,C.L. (1993) *Mol. Cell. Biol.*, **13**, 4203–4213.
6 Staples,R.R. and Dieckmann,C.L. (1993) *Genetics*, **135**, 981–991.
7 Weber,E.R. and Dieckmann,C.L. (1990) *J. Biol. Chem.*, **265**, 1594–1600.
8 Sparks,K.A., Mayer,S.A. and Dieckmann,C.L. (1997) *Mol. Cell. Biol.*, **17**, 4199–4207.
9 Mayer,S.A. and Dieckmann,C.L. (1991) *Mol. Cell. Biol.*, **11**, 813–821.
10 Sparks,K.A. (1997) PhD thesis, University of Arizona.
11 Eick,D., Wedel,A. and Heumann,H. (1994) *Trends Genet.*, **10**, 292–296.
12 St John,T.P. and Davis,R.W. (1981) *J. Mol. Biol.*, **152**, 285–315.
13 Buckholz,R.G. and Cooper,T.G. (1983) *Mol. Cell. Biol.*, **3**, 1889–1897.
14 Kukuruzinska,M.A. and Robbins,P.W. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 2145–2149.
15 Elble,R. and Tye,B.-K. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10966–10970.
16 Passmore,S., Maine,G.T., Elble,R., Christ,C. and Tye,B.-K. (1988) *J. Mol. Biol.*, **204**, 593–606.
17 Najarian,D., Dihanich,M.E., Martin,N.C. and Hopper,A.K. (1987) *Mol. Cell. Biol.*, **7**, 185–191.
18 Minvielle-Sebastia,L., Winsor,B., Bonneaud,N. and Lacroute,F. (1991) *Mol. Cell. Biol.*, **11**, 3075–3087.
19 Stone,E.M., Swanson,M.J., Romeo,A.M., Hicks,J.B. and Sternglanz,R. (1991) *Mol. Cell. Biol.*, **11**, 2253–2262.
20 Finnegan,P.M., Payne,M.J., Keramidaris,E. and Lukins,H.B. (1991) *Curr. Genet.*, **20**, 53–61.
21 Minvielle-Sebastia,L., Preker,P.J. and Keller,W. (1994) *Science*, **266**, 1702–1705.
22 Koerner,T.J., Myers,A.M., Lee,S. and Tzagoloff,A. (1987) *J. Biol. Chem.*, **262**, 3690–3696.
23 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
24 Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.
25 Hanahan,D. (1985) In Glover,D.M. (ed.), *DNA Cloning: A Practical Approach*. IRL Press, Oxford, UK, Vol. I, pp. 109–135.
26 Guo,Z. and Sherman,F. (1995) *Mol. Cell. Biol.*, **15**, 5983–5990.

27 Russo,P., Li,W.-Z., Guo,Z. and Sherman,F. (1993) *Mol. Cell. Biol*., **13**, 7836–7849.
28 Heidmann,S., Obermaier,B., Vogel,K. and Domdey,H. (1992) *Mol. Cell. Biol*., **12**, 4215–4229.
29 Sadhale,P.P., Sapolsky,R., Davis,R.W., Butler,J.S. and Platt,T. (1991) *Nucleic Acids Res*., **19**, 3683–3688.
30 Peterson,J.A. and Myers,A.M. (1993) *Nucleic Acids Res*., **21**, 5500–5508.
31 Egli,C.M., Springer,C. and Braus,G.H. (1995) *Mol. Cell. Biol*., **15**, 2466–2473.
32 Heidmann,S., Schindewolf,C., Stumpf,G. and Domdey,H. (1994) *Mol. Cell. Biol*., **14**, 4633–4642.
33 Manley,J.L. and Takagaki,Y. (1996) *Science*, **274**, 1481–1482.
34 Guo,Z. and Sherman,F. (1996) *Trends Biochem. Sci*., **21**, 477–481.
35 Guo,Z. and Sherman,F. (1996) *Mol. Cell. Biol*., **16**, 2772–2776.
36 Irniger,S. and Braus,G.H. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 257–261.
37 Nevins,J.R. and Darnell,J.E.,Jr. (1978) *Cell*, **15**, 1477–1493.
38 Mann,K.P., Weiss,E.A. and Nevins,J.R. (1993) *Mol. Cell. Biol*., **13**, 2411–2419.
39 MacDonald,C.C., Wilusz,J. and Shenk,T. (1994) *Mol. Cell. Biol*., **14**, 6647–6654.
40 Takagaki,Y. and Manley,J.L. (1997) *Mol. Cell. Biol*., **17**, 3907–3914.
41 Edwalds-Gilbert,G. and Milcarek,C. (1995) *Mol. Cell. Biol*., **15**, 6420–6429.
42 Takagaki,Y., Seipelt,R.L., Peterson,M.L. and Manley,J.L. (1996) *Cell*, **87**, 941–952.
43 Kessler,M.M., Zhao,J. and Moore,C.L. (1996) *J. Biol. Chem*., **271**, 27167–27175.
44 Amrani,N., Minet,M., Wyers,F., Dufour,M.-E., Aggerbeck,L.P. and Lacroute,F. (1997) *Mol. Cell. Biol*., **17**, 1102–1109.
45 Minvielle-Sebastia,L., Preker,P.J., Wiederkehr,T., Strahm,Y. and Keller,W. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 7897–7902.
46 Mandart,E. and Parker,R. (1995) *Mol. Cell. Biol*., **15**, 6979–6986.
47 Takagaki,Y. and Manley,J.L. (1994) *Nature*, **372**, 471–474.
48 Kessler,M.M., Henry,M.F., Shen,E., Zhao,J., Gross,S., Silver,P.A. and Moore,C.L. (1997) *Genes Dev*., **11**, 2545–2556.
49 Bolivar,F., Rodriguez,R.L., Greene,P.J., Betlach,M.C., Heyneker,H.L., Boyer,H.W., Crosa,J.H. and Falkow,S. (1977) *Gene*, **2**, 95–113.
50 Dieckmann,C.L., Homison,G. and Tzagoloff,A. (1984) *J. Biol. Chem*., **259**, 4732–4738.
51 Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.