

Signal sequence for generation of mRNA 3' end in the *Saccharomyces cerevisiae* GAL7 gene

A.Abe¹, Y.Hiraoka² and T.Fukasawa^{2,3,4}

¹The Kitasato Institute, 5-9-1, Shirokane, Minato-ku, Tokyo 108, Japan, ²Department of Microbiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160, Japan, and ³Laboratory of Molecular Genetics, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160, Japan

⁴Corresponding author

Communicated by W.Keller

We have identified a signal sequence (designated core signal) necessary to specify formation of mRNA 3' end of the GAL7 gene in *Saccharomyces cerevisiae* within a DNA segment 26 bp long. The sequence was located 4–5 nucleotides upstream from the 3' end, i.e. the polyadenylation site, of the GAL7 mRNA. Replacement of a DNA segment encompassing the polyadenylation site with a pBR322 DNA, leaving the core signal intact, resulted in alteration of the mRNA 3' end by several nucleotides, suggesting the existence of an additional signal (designated end signal) at or near the polyadenylation site. The normal end formation was abolished when the core signal was placed in the reverse orientation. A considerable fraction of pre-mRNA synthesized *in vitro* with SP6 RNA polymerase on the template of a DNA fragment containing these signals was cleaved and polyadenylated presumably at the *in vitro* 3' end during incubation in a cell-free system of yeast. By contrast pre-mRNA synthesized on the template with the core signal alone was processed but much less efficiently. No such processing was seen when the pre-mRNA either lacked the core signal or contained it in the reverse orientation.
Key words: deletion analysis/*in vitro* processing/polyadenylation site/yeast

Introduction

Proper formation of mRNA 3' end is essential for the optimal expression of a gene. The mechanism by which mRNA 3' ends are formed has been studied extensively in both prokaryote and higher eukaryote. Now we know that in the majority of prokaryotic genes, transcription termination leads directly to the 3' end formation of mRNA (Platt, 1986). This contrasts with the mechanism in higher eukaryote, where formation of mRNA 3' ends is a complex process that involves a nucleolytic cleavage of long primary transcripts at specific sites followed by polyadenylation of newly generated ends (Proudfoot and Whitelaw, 1988). In the yeast *Saccharomyces cerevisiae*, however, two lines of experiments are apparently contradictory to each other: (i) Zaret and Sherman suggested that the 3' end of *CYC1* (iso-1-cytochrome *c*) mRNA is formed by direct polyadenylation of the 3' end of primary transcript (Zaret and Sherman, 1982). They proposed TAG ... TAGT or TATGT...(AT rich)...TTT as a consensus signal for

transcription termination in many genes including *CYC1* (Zaret and Sherman, 1982). Deletion of a 38 bp sequence in the signal results in the generation of various mRNAs with longer than normal sizes, all of which are polyadenylated at the respective 3' ends. This model has received further support from Osborne and Guarente (1989), who showed, by a combination of mutational analysis and transcriptional run-on experiments, that an 83 bp sequence in between convergently transcribed genes, *CYC1* and *UTR1* (unidentified transcript), is sufficient for generation of the 3' ends of both RNAs and that the sequence contains two transcriptional terminators for the respective genes. (ii) On the other hand, Butler and Platt (1988) demonstrated that pre-mRNA synthesized *in vitro* with SP6 RNA polymerase on the template of a DNA fragment encompassing the intergenic 237 bp sequence between *CYC1* and *UTR1* is cleaved at, or very close to the *in vivo* polyadenylation site during incubation in a whole cell extract of yeast in the presence of ATP. They further showed that one of the cleaved products is polyadenylated in the reaction mixture. Most recently they partially purified a subcellular fraction containing the processing activity from the whole cell extract, which accurately and efficiently cleaves and polyadenylates a variety of yeast pre-mRNAs (Butler *et al.*, 1990). These include those of *CYC1*, *HIS4* (histidine biosynthesis), *GAL7* (galactose-1-phosphate uridylyltransferase), *H2B2* (histone H2B2), *PRT2* (a protein of unknown function) and *CBP1* (cytochrome *b* mRNA processing). These results have strongly suggested that formation of mRNA 3' ends in yeast is analogous to that in higher eukaryotes.

In the past several years, we have studied sequence elements in the *GAL7* gene necessary for formation of 3' end of its mRNA by mutational analysis as well as by a series of experiments with synthetic DNA fragments that mimic the sequence. Recently we have found the necessary sequence to be located 4–5 nucleotides upstream from the polyadenylation site within a discrete region 26 bp long. The sequence functions normally only in the right orientation. Furthermore, a region encompassing the polyadenylation site may contain an auxiliary signal to determine the mRNA 3' end, since replacement of that region with a pBR322 DNA leads to alteration of the 3' end by several nucleotides without affecting the efficiency of its formation. We have also studied whether or not *in vitro* synthesized pre-mRNA containing this determined sequence is specifically cleaved and polyadenylated during incubation in a partially purified fraction of yeast cell extract prepared according to the procedures of Butler *et al.* (1990). Cleavage and adenylation occurred in a significant fraction of pre-mRNA synthesized *in vitro* at or near a site corresponding to the *in vivo* polyadenylation site, when the pre-mRNA contained both core and end signals. Taking these findings into account, we suggest that the sequences we identified constitute signals that are necessary for formation of the 3' end of *GAL7*

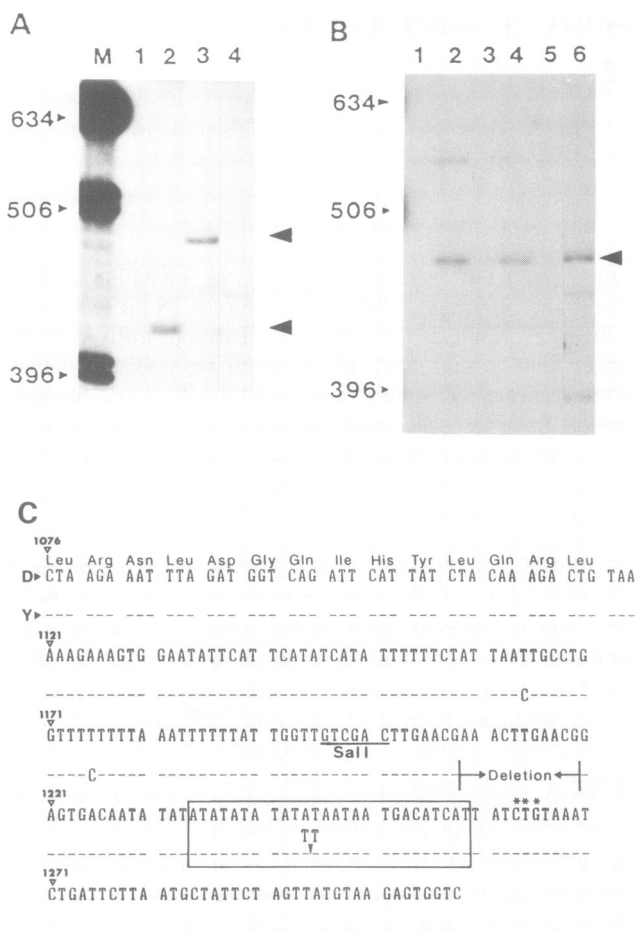


Fig. 1. Determination of 3' end of *GAL7* mRNA in two yeast strains. (A) S1-protected fragments obtained by the use of a probe DNA derived from strain D585-11C (the *Hind*III–*Xba*I fragment of pYH3002). Lanes 2 and 3 contained mRNA from strain YK3 and from D585-11C, respectively; both strains were grown in the presence of galactose. Lane 1 contained no mRNA, whereas lane 4 contained mRNA from strain D585-11C grown in the absence of galactose. Size markers (lane M) were 5' end-labeled fragments of pBR322 digested with *Eco*RI and *Hin*FI. The arrowheads on the right indicate the major S1-protected bands. (B) S1-protected fragments obtained by use of a probe DNA derived from strain YK3 (the *Hind*III–*Xba*I fragment of pRGY7). Lanes 2, 4 and 6 contained mRNA from *gal7* null yeast (N590) carrying a plasmid with *GAL7* whose 3' flanking region was derived from YK3, mRNA from YK3 itself and mRNA from strain N590 carrying a plasmid with *GAL7* derived from D585-11C, respectively. All these yeasts were grown in the presence of galactose. Lane 1 contained no RNA whereas lanes 3 or 5 mRNA from the same yeasts in lanes 2 or 4 but grown in the absence of galactose. Size markers were the same as used in (A). The arrowheads on the right indicate the major S1-protected bands. (C) Nucleotide sequence of a 3' region of *GAL7* in strains D585-11C (upper line; Tajima *et al.*, 1985) or YK3 (lower line; this work). Dashed lines indicate identical bases between two strains. The figures at the beginning of each line signify the number of bases starting from the transcription initiation site of *GAL7*. The boxed region is an essential sequence identified by a series of deletion analyses in the following experiments. Probable 3' ends are marked with asterisks.

mRNA *in vivo*, and that processing of a long pre-mRNA is involved in the 3' end formation.

Results

Determination of 3' end of *GAL7* mRNA in different yeast strains

Previously we determined 3' end of *GAL7* mRNA in a yeast strain, YK3, by the S1 protection mapping using a DNA

fragment of *GAL7* derived from another strain, D585-11C as the probe (Tajima *et al.*, 1985). To our surprise, when we similarly determined the 3' end of *GAL7* mRNA in strain D585-11C, the size of S1-protected fragments was different from that in YK3; 477 nucleotides (nt) for D585-11C and 417 nt for YK3 (Figure 1A). The result could occur for either of the following reasons. (i) The 3' ends of *GAL7* mRNAs in these strains were located at different positions. (ii) There was a small deletion in the 3' untranslated region of *GAL7* mRNA in YK3, which caused formation of an S1-protected fragment shorter than expected. We therefore retrieved a DNA segment of 3' flanking region from YK3 to use as a source of probe DNA. When both mRNA and probe DNA were prepared from YK3, we found a protected fragment with a size identical with that observed when both mRNA and probe DNA were derived from D585-11C (Figure 1B); the result agreed with the second model. Indeed, sequencing of the retrieved DNA fragment revealed an 11 bp deletion which was located at a site expected from the size of the protected fragment in Figure 1A. In addition, two point mutations and a two-base insertion were seen in YK3 (Figure 1C). The insertion was located within a region that turned out to contain an essential signal for the 3' end formation of *GAL7* mRNA in the following experiments.

Localization of signal sequence for mRNA 3' end formation in *GAL7*

Recently we found that a sequence necessary for generating the normal 3' end of *GAL7* mRNA was located within 328 bp fragment between the *Sall* and *Xba*I sites in 3' flanking region of *GAL7* (Abe *et al.*, 1988; see Figure 6). To determine more precisely the 3' margin of the sequence that signals mRNA 3' end formation, we constructed a series of deletions extending upstream from the *Xba*I site. Those deletions (Figure 2A, 1–6) were inserted in a multi-copy plasmid (pNI2) and introduced to a *gal7* null yeast (N590). Total RNA was extracted from the transformant yeasts grown in galactose-containing medium and subjected to Northern analysis. As is clearly seen in Figure 2B (lanes 1–6), mRNA of a discrete size was produced if a plasmid retained >58 bp downstream from the *Sall* site. By contrast, yeast with a plasmid lacking a region downstream of the *Sall* site produced mRNA molecules with various sizes, most of which were larger than the normal mRNA. A quantitatively similar result was obtained if poly(A)⁺ RNA was used for analysis (data not shown), indicating that practically all the abnormal mRNA molecules were polyadenylated as in the case of *cyc1-512*, a mutant yeast which was originally suggested to be defective in the transcriptional terminator of the *CYC1* gene (Zaret and Sherman, 1982). To determine the 5' margin of the signal sequence, we constructed a series of deletions, in which sequence downstream from the *Sall* site in the 3' flanking region of *GAL7* was deleted to various extents (Figure 2A, 7–11). Each of the *GAL7* genes bearing those deletions was carried by pNI2 and introduced into the *gal7* null yeast. Northern analysis of total RNA from the respective transformants clearly indicated that the 5' end of the signal was located between 33 bp and 43 bp downstream from the *Sall* site (Figure 2B; lanes 7–11). Deletions whose 3' ends were located <33 bp downstream from the *Sall* site (Figure 2A, 7–9) were able to form a normal 3' end of the mRNA. In contrast, in deletions that extended >43 bp, most of the transcripts exhibited various sizes <1.3 kb, the

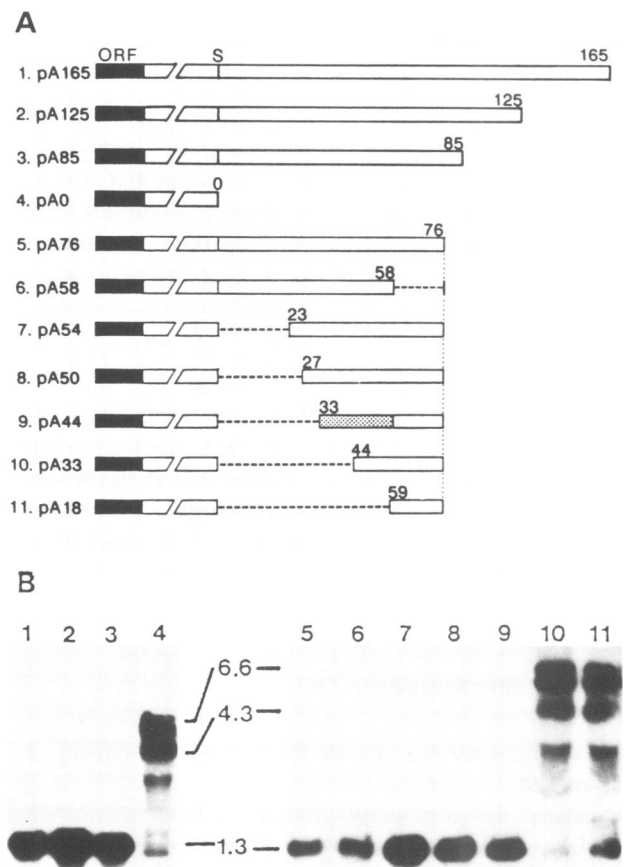


Fig. 2. Determination of the ends of the signal sequence specifying the 3' end of *GAL7* mRNA. (A) Structure of *GAL7* bearing various deletions. Boxes represent yeast DNA, in which solid and stippled area indicate the coding region of *GAL7* and 'core signal' (see the text), respectively. Designations on the left of the figures represent plasmids that carry the respective deletions. The distance between the boundary of a deletion and the *SalI* site (S) is shown above the boxes in base pairs. (B) Northern analysis of *GAL7* mRNA produced by *gal7* null yeast carrying the indicated plasmids. The electrophoreses for lanes 1–4 and 5–11 were carried out separately. The sizes of major species of mRNA were indicated between lanes 4 and 5, which are extrapolations from the positions of large (3.2 knt) and small (1.6 knt) rRNAs.

size of the normal *GAL7* mRNA. Taking these and the above results into account, we concluded that the main signal necessary for mRNA 3' end formation was located within the 26 bp fragment between nucleotide positions 33 bp and 58 bp (Figure 2A). This signal will hereafter be referred to as 'core signal'. Note that a small fraction of the transcripts in lanes 4, 10 and 11 exhibited a size slightly shorter than 1.3 nt. The significance of these transcripts will be discussed later.

Probable presence of additional signal element at or near the poly(A) site

In the above experiment, the polyadenylation site itself appeared to be removable without affecting the efficiency of mRNA 3' end formation (see Figure 2A, pA58). However, there still remained the possibility that an additional element existed at or near the polyadenylation site necessary for determining the 3' end, since the size determination of mRNA by Northern analysis in agarose gel is not sufficiently precise. We therefore determined the 3' end by the RNase A protection method in the transcript from

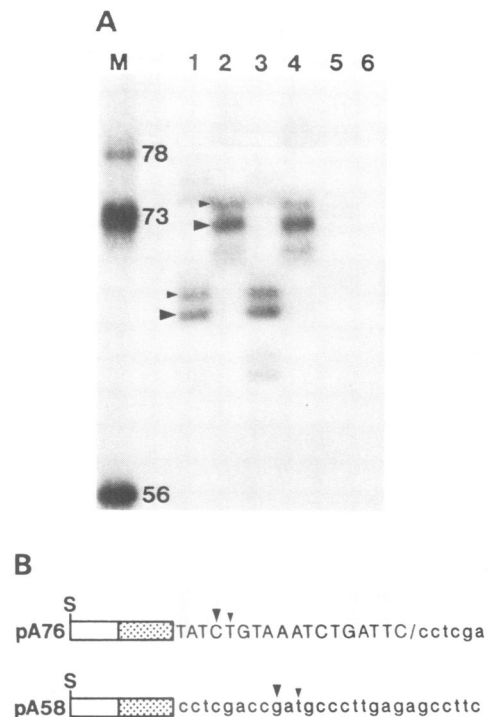


Fig. 3. Determination of 3' ends of *GAL7* mRNA directed by two plasmids by RNase A protection mapping. (A) Total RNA from a *gal7* null yeast bearing either pA76 (lanes 1 and 3) or pA58 (lanes 2 and 4) grown in galactose-containing medium were hybridized with 32 P-labeled probe RNA and digested with 3 μ g (lanes 1 and 2) or 6 μ g (lanes 3 and 4) of RNase A. Lanes 5 and 6 contained mRNAs from the same yeasts in lanes 3 and 4, respectively, but grown in the absence of galactose. Lane M contained size markers of the indicated sizes (in nucleotides). The large and small arrowheads indicate the major and minor RNase A-protected bands. (B) Nucleotide sequences encompassing the 3' ends determined in (A). Clear and dotted boxes represent the *GAL7* 3' flanking region and the 'core signal' (see the text), respectively. Capital and lower case letters represent sequences of yeast and pBR322 DNA, respectively. Large and small arrowheads respectively correspond to the major and minor protected bands in (A). S represents the *SalI* site.

plasmid pA76 bearing the entire 3' transcribed region of *GAL7* or from pA58 lacking a sequence encompassing the polyadenylation site by replacement with pBR322 DNA. As is seen in Figure 3A, the 3' end of the respective transcripts differed by 4–5 nt, suggesting the presence of an additional element at or near the 3' end of *GAL7* mRNA, which was involved in some way in precise determination of the 3' end. Since extensive homology was not apparent between *GAL7* and pBR322 sequences that were contained in the respective plasmids (see Figure 3B), the element in question may consist of a few bases. This presumptive element will be designated 'end signal' from now on.

Effect of synthetic oligonucleotides that mimic signal sequence on 3' end formation

A remarkable feature in the 26 bp sequence identified in the above experiments is an AT alternating stretch (see Figure 2C). To know whether or not the AT alternating sequence was sufficient for the 3' end formation, we synthesized an oligonucleotide of (AT)₇ with *SalI* linkers at both ends, inserted it at the *SalI* site in 3' flanking region of *GAL7* (Figure 4A), and studied the effect on the formation of mRNA end. None of the insertions led to production of *GAL7* mRNA with the authentic size, indicating that

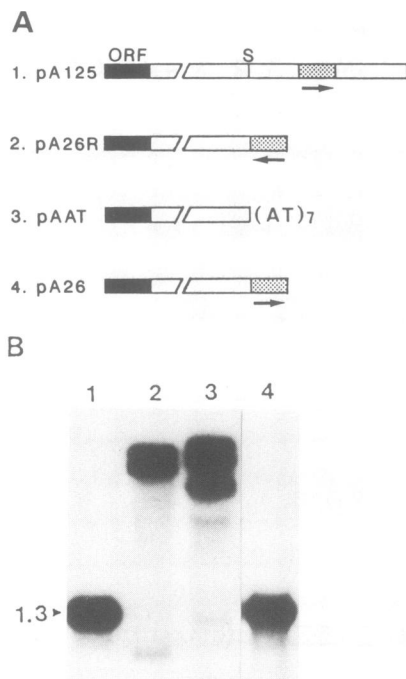


Fig. 4. Effect of synthetic oligonucleotides on the 3' end formation of *GAL7* mRNA. (A) Structure of *GAL7* bearing synthetic sequences that mimic the 'core signal' for 3' end formation in its 3' flanking region. Plasmid pA26R lacks the *Bgl*III–*Hind*III fragment of 0.4 kb in the *GAL7* coding region. Solid, clear and dotted boxes represent coding region, non-translatable region and the 'core signal' (see the text). The arrows indicate directions of the core signal. (B) Northern blot hybridization of mRNA directed by *GAL7* bearing synthetic sequences that mimic the 'core signal' for 3' end formation. S represents the *Sal*I site.

(AT)₇ was incapable of signaling the 3' end formation (Figure 4B; pAAT). We then synthesized an oligonucleotide with a sequence of (AT)₆AATAATGACATCAT that mimics entirely the presumed signal sequence, inserted it at the *Sal*I site in the 3' flanking region of *GAL7* through the *Sal*I linker at the 5' end and *Xho*I linker at the 3' end, and studied its effect on the size of *GAL7* mRNA. As is clearly seen in Figure 4B, mRNA with a discrete size was synthesized in yeast carrying a plasmid containing the oligonucleotide in the right orientation. The size coincided with the position of the insert. When the synthetic 26 bp fragment was inserted in the reverse orientation at the *Sal*I site, mRNAs with various sizes were yielded, most of which were longer than that found in the plasmid with the sequence in the right orientation. These results suggest that the whole 26 bp sequence of the core signal was necessary for the formation of mRNA 3' end *in vivo*, and that the sequence functions only in the right orientation.

In vitro processing of pre-mRNA of *GAL7*

We constructed a plasmid bearing a 461 bp fragment of *GAL7* DNA containing the core signal as well as the polyadenylation site (Figure 5A, pSP44). Using the plasmid linearized at the *Pvu*II site, capped RNA of 692 nt was synthesized with SP6 RNA polymerase *in vitro* in the presence of [α -³²P]UTP. The transcript was incubated in a partially purified fraction from whole yeast cell extract at 30°C for the indicated times. Two species of RNA (520 nt and 430 nt) were found to be produced during the incubation,

when the whole reaction mixtures were fractionated by electrophoresis in urea–polyacrylamide gel (Figure 5B). We then selected polyadenylated RNA from the 60 min sample with oligo(dT) latex beads and electrophoresed along with the original sample. As seen in the same figure, both RNA products were polyadenylated. The size of the longer product was successfully accounted for by assuming that cleavage and polyadenylation with ~60 nucleotide residues occurred at the *in vivo* polyadenylation site. To study requirements for cofactors for the *in vitro* processing reaction, various nucleoside triphosphates were incorporated in reaction mixtures in place of ATP. As shown in Figure 5C, the nucleoside triphosphates other than ATP led to production of neither species of poly(A)⁺ RNA, indicating that production of both poly(A)⁺ RNAs was an ATP dependent reaction. Note that a species of RNA shorter than the poly(A)⁺ RNA of 520 nt was synthesized in the reaction mixtures containing dATP, UTP, cordycepin or the mixture without cofactor. The size of this RNA product was estimated to be ~460 nt, which coincided with that of non-poly(A) part of the poly(A)⁺ RNA (see Figure 5E below), suggesting that only cleavage but not polyadenylation occurred to the pre-mRNA when a reaction mixture contained no exogenous ATP.

Requirement for the signals for *in vitro* processing of pre-mRNA

To study the mechanism for the occurrence of two poly(A)⁺ RNAs, we prepared pre-mRNAs on various templates which carried alterations in the signal elements (Figure 5A): pSP26 and pSP26R, lacking a fragment encompassing the polyadenylation site, contained the synthetic core signal in the right and reverse orientations, respectively. Plasmid pSPDT was devoid of the entire region downstream the *Sal*I of the *GAL7* DNA in pSP44, whereas pSP33 lacked the AT alternating sequence in the core signal, leaving the other *GAL7* region intact. When these pre-mRNAs were incubated in the processing mixture as above, a small amount of the 520 nt poly(A)⁺ RNA was produced only from the pSP26 pre-mRNA (Figure 5D). In contrast, the 430 nt poly(A)⁺ RNA was yielded from all the pre-mRNAs. From these results, the following conclusions may be drawn. (i) Production of the 520 nt poly(A)⁺ RNA absolutely requires for the presence of the core signal in the right orientation. (ii) Occurrence of 430 nt poly(A)⁺ RNA is independent of both signals, and possibly directed by an unknown element located upstream from the *Sal*I site. To estimate the length of poly(A) stretches in the *in vitro* products, aliquots of the poly(A)⁺ RNA samples were hybridized with oligo(dT)_{12–18} and digested with RNase H. The resulting non-poly(A) RNAs were separated by electrophoresis along with the respective poly(A)⁺ RNAs. The result in Figure 5E suggests that both 520 nt and 430 nt poly(A)⁺ RNAs were shortened by ~60 nt. We estimate therefore ~60 adenosine residues to be attached to 3' ends of the *in vitro* products.

Discussion

In the present work, we located a major (core) signal for the 3' end formation of the *GAL7* mRNA within a 26 bp segment of DNA in its 3' flanking region. Within the sequence, a two-base insertion was found in one of the two

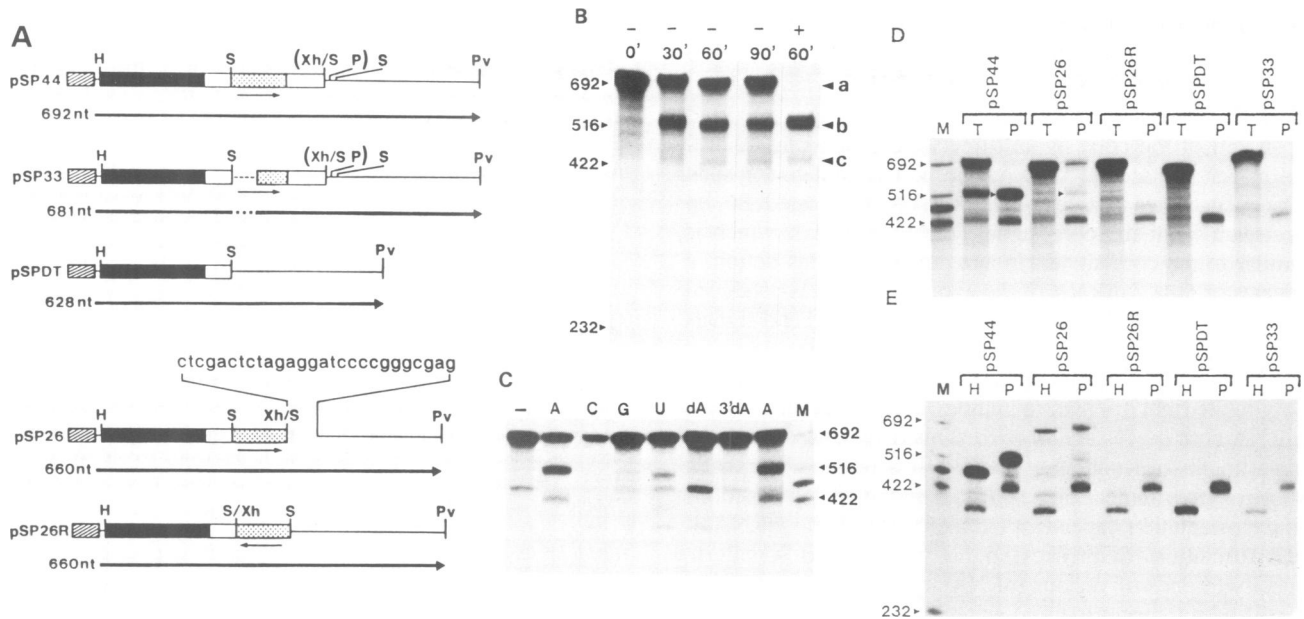


Fig. 5. *In vitro* processing of pre-mRNA of *GAL7*. (A) Constructions of plasmids used for preparing various pre-mRNA. Stippled, solid, dotted and clear boxes represent the SP6 promoter, *GAL7* coding region, the 'core signal' (see the text) and the *GAL7* 3' flanking region, respectively. Thin and broken lines represent pSP64 DNA and a deleted region in the 'core signal', respectively. Restriction sites are H, *Hind*III; S, *Sal*I; Xh, *Xho*I; P, *Pst*I and Pv, *Pvu*II. A short segment containing the *Pst*I and *Sal*I sites in the parentheses in pSP44 and pSP33 was derived from pUC118. Thin (short) and thick (long) arrows under the plasmids are directions of the 'core signal' and pre-mRNAs, respectively. A sequence immediately downstream of the core signal in pSP26 is shown above its restriction map for comparison with the pBR322 sequence in Figure 3B (see Discussion). (B) Autoradiogram of the gel electrophoretic separation of the products of the incubation of *GAL7* pre-mRNA synthesized on the template of pSP44 in (A). The incubation times in minutes are shown across the top. Symbols + or - above the incubation times indicate respectively that the samples were or were not subjected to the selection of poly(A)⁺ RNA. The positions of size markers are shown in nucleotides at the left. Arrowheads a, b and c indicate the pre-mRNA, the major and minor products, respectively. (C) Requirement for cofactors for the *in vitro* processing of *GAL7* pre-mRNA. The *GAL7* pre-mRNA synthesized on the template of pSP44 (see above) was incubated in the presence of various nucleoside triphosphates at a final concentration of 2 mM. The products were separated by electrophoresis and autoradiographed. Lanes (-), A, C, G, U, dA, and 3'dA represent reaction mixtures containing no nucleotide cofactors, ATP, CTP, GTP, UTP, deoxy-ATP and cordycepin, respectively. Mixtures of the processing reaction were incubated for 60 min at 30°C. Lane M contained size markers, whose lengths are indicated to the right. (D) Selection of poly(A)⁺ RNA from *in vitro* processed products of various *GAL7* pre-mRNAs. Template plasmids, on which the *GAL7* pre-mRNAs were synthesized are shown at the top. Lanes T and P represent total products and poly(A)⁺ RNA, respectively. The arrowheads in the autoradiogram indicate core signal dependent RNA products. Mixtures of the processing reaction were incubated for 60 min at 30°C. Lane M contained size markers, whose lengths are shown at the left. (E) Estimation of the length of poly(A) stretches in poly(A)⁺ RNA products from various *GAL7* pre-mRNA. Template plasmids, on which the *GAL7* pre-mRNAs were synthesized are shown at the top. Mixtures of the processing reaction were incubated for 60 min at 30°C. Lanes H and P contained poly(A) deleted and poly(A) attached RNA samples, respectively.

strains studied, where the 3' end of *GAL7* mRNA was identical. This result indicated that strict sequence specificity is not required for the signal function. We also suggested existence of additional (end) signal at or near the polyadenylation site based on the finding that the region encompassing the presumptive end signal, when replaced with an arbitrary DNA segment that has no extensive homology, resulted in an alteration of the polyadenylation site. The replacement, however, caused no appreciable change in the efficiency of the end generation *in vivo*.

We further studied the role of the signal elements in the *in vitro* processing of pre-mRNA containing the signal sequences using the cell-free system of yeast prepared by the procedures described by Butler *et al.* (1990). These results demonstrated that the core signal was essential for specific cleavage as well as polyadenylation of pre-mRNA. When the sequence encompassing the presumptive end signal was replaced with a pSP64 sequence, the processing occurred but much less efficiently. This contrasts with the *in vivo* experiment, in which replacement of the sequence containing the presumptive end signal with a pBR322 was shown to result in an alteration of the polyadenylation site by several nucleotides, but not appreciably in the efficiency of its generation. We assume that the difference between the pSP64

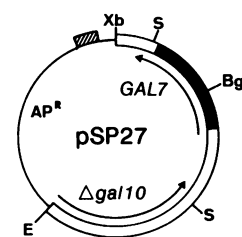


Fig. 6. Construction of plasmid pSP27 which carries the *GAL7* gene. Solid and clear boxes represent *GAL7* coding and flanking regions, respectively. The plasmid also contains *GAL10* whose 5' region is deleted. Arrows inside the plasmid indicate the direction of transcription of the genes. The thin lines, stippled box downstream of *GAL7* and Ap^R represent pBR322 DNA, the SP6 promoter and the ampicillin resistance gene, respectively. The restriction sites are Xb, *Xba*I; S, *Sal*I; Bg, *Bgl*II and E, *Eco*RI.

(Figure 5A; pSP26) and pBR322 (Figure 3B; pA58) sequences caused the discrepancy between the *in vivo* and *in vitro* experiments. Examination of such a possibility may give us further insight into the nature of the end signal. In addition to the RNA product with the authentic 3' end, our *in vitro* reaction yielded a species of poly(A)⁺ RNA, whose

generation was independent of the core signal. We believe that this product was not a mere artifact of the *in vitro* reaction, since the generation of this RNA required ATP. Instead we assume that yet unidentified signal located upstream of the core signal directs the processing of pre-mRNA at the second site. As the best candidate for such a sequence, we noticed CATTTCATATC located 86 bp upstream from the core signal (see Figure 1C), which is similar to the core signal. In agreement with this notion, a species of *GAL7* mRNA of the expected size was produced in a small amount *in vivo*, even if both core and end signals were deleted (Figure 2B; lanes 4, 10 and 11).

We have also demonstrated that the core signal functions only in the right orientation in the 3' end formation of *GAL7* mRNA both *in vivo* and *in vitro*. This result agrees with the experiments of Ruohola *et al.* (1988) which indicate that a short DNA fragment of *CYCI* 3' flanking region functions in the orientation dependent manner to direct the mRNA 3' end formation. Sequences so far proposed as signals to specify 3' ends of various mRNAs in yeast (Yu and Elder, 1989 and references therein) exhibit no extensive similarities with each other except that some of them contain the consensus sequence proposed by Zaret and Sherman (1982). One might argue that no single mechanism is responsible for mRNA 3' end formation. On the contrary, however, the fact that a variety of pre-mRNAs are processed to yield mature mRNAs in the same cell-free system would strongly suggest that a common mechanism underlies this processing (Butler *et al.* 1990). Recently an snRNA has been suggested to be involved in 3' end formation of histone mRNA in sea urchin as well as in mouse (Birnstiel *et al.*, 1985; Cotten *et al.*, 1985; Gick *et al.*, 1986). By an analogy, it is tempting to speculate that snRNA plays a crucial role in the 3' end formation also in yeast. This hypothesis assumes base pairings between pre-mRNA and snRNA, allowing non-matching spacers in between the paired regions in the 'processing complex'. Such a hypothesis would successfully explain the observed absence of a highly conserved nucleotide sequence among the proposed signals and also their dipartite (*GAL7*; present work) or tripartite (*CYCI*; Zaret and Sherman, 1982) nature. A variety of snRNAs may be assumed to be responsible for the formation of mature 3' ends of a variety of mRNAs in yeast.

Materials and methods

Strains and media

S.cerevisiae strains were N590 (*MATa ade ura3 leu2 trp1 his3 gal7Δ::LEU2*), YK3 (*MATa trp1 his3*), YA4 (*MATa trp1 ura3::HIS3*), D585-11C (*MATa MAL2 lys2*) and TSPEP4 (*MATa ade2 ura3 trp1 his3 pep4-3*). *Escherichia coli* strains were M15 (*d[lac-pro] thiφ80dlacZM15 ara recA rpsL*) or HB101 (*supE44 hsdS20 recA ara pro lac gal rpsL xyl mut*). Yeast media, SGlu, SGlyLac, and SGal contained 0.66% yeast nitrogen base, 0.5% casamino acids, 0.002% L-tryptophan, 0.002% adenosine sulfate and 2% of each of the following carbon sources, respectively; glucose, glycerol and sodium lactate, and galactose.

Plasmids

pN12 is a derivative of YEp24 (Botstein *et al.*, 1979), in which the original *EcoRI* sites were eliminated, and the *BamHI* site was converted to *EcoRI*. pNA3 was constructed from pN12 by eliminating the *XbaI* site. pSP27 (Figure 6) contained an *EcoRI*-*XbaI* fragment of 4.2 kb in length encompassing the entire *GAL7* gene and 3' part of the *GAL10* gene, which was derived from pYH3002 (Hashimoto *et al.*, 1983); the fragment was originated from *λgt10Sc481* (St John and Davis, 1979).

Enzymes

Restriction enzymes, Bal31 exonuclease, exonuclease III (ExoIII), S1 nuclease, Klenow fragment of *E.coli* DNA polymerase, RNase H and T4 DNA ligase were purchased from Takara Shuzo Co. RNase A was a product of Sigma Chemical Co.

Northern blot analysis

Isolation of RNA. Cells of a *gal7* null yeast (N590) carrying a plasmid with various derivatives of *GAL7* were grown to mid-log phase in SGlyLac to early log phase. Galactose was added to a final concentration of 2% when the density reached $2-4 \times 10^6$ cells/ml, and the cells were collected after 3 h. Total RNAs were extracted from these cells by a rapid method (Elder *et al.*, 1983). Poly(A)⁺ RNA was selected by oligo(dT)-latex (Takara Shuzo Co.).

Preparation of probe RNA. *GAL7* anti-sense RNA for its coding region was synthesized *in vitro* in the presence of [α -³²P]UTP by the SP6 system (Amersham kit RPN 1506) on the template of the 1.3 kb DNA fragment between *Sall* and *AluI* of the *GAL7* gene in pYH3003 (Hashimoto *et al.*, 1983). The latter site was converted to *BamHI* site upon insertion into pSP64 (Melton *et al.*, 1984).

Blot hybridization. This was carried out by the published procedure (Thomas, 1980) using a Biodyne membrane (Pall Corp).

Retrieving a DNA segment from 3' flanking region of GAL7

This was carried out essentially by the method described by Hicks *et al.* (1982). The *EcoRI* fragment encompassing the entire *GAL7* gene from pYH3002 was inserted in pNA3 at the *EcoRI* site. The resulting plasmid was digested with *XbaI* and *BglIII* to eliminate a fragment of 1.1 kb containing 3' part of the coding region and its contiguous flanking region. A yeast strain YA4, a *ura3::HIS3* derivative of YK3, was transformed with the linearized plasmid, and Ura⁺ transformants were obtained on uracil-lacking SGlu agar. Plasmids were recovered from several transformants, and desired plasmids were screened by restriction mapping. The *DdeI*-*AvaII* fragment of 0.2 kb was excised from one of those plasmids (pRGY7), blunt-ended with Klenow fragment, and inserted in M13mp19 in either orientation at the *SmaI* site for sequencing.

Determination of mRNA 3' ends

S1 nuclease protection. The 3' terminus of the *GAL7* mRNA was determined by the S1 nuclease method (Weaver and Weissmann, 1979). Total RNA was isolated from strains YK3 or D585-11C. The probe DNA was prepared by excising the 739 bp *HindIII*-*XbaI* fragment from pYH3002, following 3' end-labeling with [α -³²P]dATP at the *HindIII* site with DNA polymerase Klenow fragment.

RNase A protection. Total RNA was isolated from two *gal7* null yeast strains which carrying pA76 and pA58, respectively. The *Sall*-*PvuII* fragments of 1.5 kb were excised from both plasmids, and each fragment was inserted in pSP65 at the *SmaI* and *Sall* sites. The resulting plasmids were linearized with *Sall* and used as the template to synthesize RNA *in vitro* with SP6 RNA polymerase in the presence of [α -³²P]UTP.

Construction of 3' deletions of signal sequence

A DNA fragment of 4.2 kb encompassing the whole sequence of *GAL7* was excised from pYH3002 with *XbaI* and *EcoRI* and inserted into pSP64 at the corresponding sites to generate pSP27 (Figure 6). The plasmid was cleaved at *XbaI*, digested with Bal31, and self-ligated in the presence of *XhoI* linker following filling-in with Klenow fragment. Desired deletions were selected for size of *XhoI*-*Sall* fragments from the resulting plasmids by polyacrylamide gel electrophoresis. Ends of the deletions were determined by sequencing the *Sall*-*XhoI* fragments. Five plasmids with deletions ranging from 163 bp to 267 bp were digested with *EcoRI* and *XhoI*, and resulting fragments with *GAL7* were inserted into pN12.

Construction of 5' deletions of signal sequence

A derivative of pSP27 was cleaved at the *Sall* site, digested with ExoIII, and self-ligated in the presence of *Sall* linker following filling-in with Klenow fragment. Desired deletions were selected for size of *Sall*-*XhoI* fragments from the resulting plasmids by polyacrylamide gel electrophoresis. Ends of the deletions were determined by sequencing of *Sall*-*XhoI* fragments. In each of five plasmids thus obtained, the *EcoRI*-*Sall* fragment containing a deletion was replaced with the *EcoRI*-*Sall* fragment from a derivative of pSP27, in which one of the two *Sall* sites in *GAL10* had been eliminated.

From the resulting plasmids, *XhoI*–*EcoRI* fragments with *GAL7* was excised and inserted into pNI2.

In vitro processing

Precursor RNA was synthesized in SP6 transcription reaction mixture composed of 40 mM Tris–HCl pH 7.5, 6 mM MgCl₂, 1 mM spermidine, 10 mM dithiothreitol (DTT), 0.5 mM each of ATP and CTP, 0.1 mM GTP, 0.02 mM UTP, 0.5 mM m⁷GpppG, 1200 µCi/ml [α -³²P]UTP, 900 U/ml ribonuclease inhibitor, 40 µg/ml of a linearized template, and 1500 units/ml SP6 RNA polymerase. Whole cell extract was prepared essentially according to the procedures established for the *in vitro* splicing in yeast (Lin *et al.*, 1985). The pre-mRNA processing activity was partially purified as described by Butler and Platt (1990) with minor modifications: cells of strain TSPEP4 were grown at 30°C to late log phase in 1 l of YPD medium. When the cell concentration reached an optical density of 4.0 at 600 nm, cells were collected and converted to spheroplasts as described by Lin *et al.* (1985). The spheroplasts were resuspended in 8 ml of buffer A [10 mM HEPES–KOH pH 7.0, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM DTT] and disrupted by vigorous agitation on a Voltex mixer for 30 s in the presence of an equal volume of glass beads with a diameter of 0.5 mm. The agitation was repeated 10 times at 1 min intervals. The sample was chilled at the intervals on ice. Solid KCl was added to the sample with stirring to give a final concentration of 0.2 M. The sample was left for another 30 min on ice and passed through a syringe with a cotton plug at the bottom to get rid of glass beads. The suspension was centrifuged at 18 000 g for 10 min in the cold, and the supernatant was further centrifuged at 146 000 g for 60 min. To the clear supernatant, solid ammonium sulfate was added with stirring to 40% saturation. Precipitated materials were collected by centrifugation at 15 000 g for 10 min in the cold, dissolved in 1 ml of buffer B [20 mM HEPES–KOH pH 7.0, 0.2 mM EDTA, 50 mM KCl, 20% (v/v) glycerol, 0.5 mM DTT] and dialyzed against 1 l of the same buffer for 3 h. The final sample was appropriately distributed into several tubes and kept frozen at –75°C until use. Processing reaction was performed at 30°C in a mixture of 10 µl, which contained 4.6 mM HEPES pH 7.0, 0.05 mM EDTA, 2 mM ATP, 1 mM magnesium acetate, 11 mM potassium chloride, 75 mM potassium acetate, 2% polyethylene glycol 6000, 0.1 mM dithiothreitol, 4% glycerol, 23% cell extract (v/v), and 10 nM RNA (20 000 c.p.m./µl). Selection of poly(A)⁺ RNA was carried out oligo(dT)–latex (Takara Shuzo. Co., Ltd) according to the supplier's protocol.

Other methods

The lithium acetate method was used for transformation of yeast (Ito *et al.*, 1983). Sequencing of DNA was performed by the dideoxynucleotide chain-termination method (Sanger, 1981) using M13 phages mp18 and mp19 (Messing, 1983). Oligonucleotides that mimic the presumed signal sequence of mRNA 3' end formation was synthesized by an automatic DNA synthesizer (Applied Biosystems). Elimination of poly(A) stretch from poly(A)⁺ RNA was essentially according to Gellersen *et al.*, (1989). Briefly, a poly(A)⁺ RNA fraction from 100 µl of the processing reaction was mixed with 0.5 µg of oligo(dT)_{12–18} (Pharmacia) in the presence of 0.1 M KCl, and the mixture was incubated successively at 78°C for 2 min and at 37°C for 30 min, followed by the treatment with 8 U of RNase H.

Acknowledgements

We are indebted to Dr Terry Platt for sending the preprint of the *in vitro* processing (Butler *et al.*, 1990) before publication. This work was supported in part by Grants-in-aids from Ministry of Education, Science and Culture, and for New Drug Development from Ministry of Health and Welfare.

References

- Abe, A., Wada, T., Handa, H., Nogi, Y. and Fukasawa, T. (1988) *Agric. Biol. Chem.*, **52**, 2035–2041.
- Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) *Cell*, **41**, 349–359.
- Botstein, D., Falco, S.C., Stewart, S.E., Brennan, M., Sherer, S., Stinchcomb, D.T., Struhl, K. and Davis, R.W. (1979) *Gene*, **8**, 17–24.
- Butler, J.S. and Platt, T. (1988) *Science*, **242**, 1270–1274.
- Butler, J.S., Sadhale, P. and Platt, T. (1990) *Mol. Cell. Biol.*, **10**, 2599–2605.
- Cotten, M., Gick, O., Vasserot, A., Schaffner, G. and Birnstiel, M. (1988) *EMBO J.*, **7**, 801–808.
- Elder, R.T., Loh, E.Y. and Davis, R.W. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2432–2436.
- Gellersen, B., DiMattia, G.E., Friesen, H.G. and Bohnet, H.G. (1989) *Mol. Cell. Endocrinol.*, **64**, 127–130.
- Gick, O., Keller, W. and Birnstiel, M.L. (1986) *EMBO J.*, **5**, 1319–1326.
- Hashimoto, H., Kikuchi, Y., Nogi, Y. and Fukasawa, T. (1983) *Mol. Gen. Genet.*, **191**, 31–38.
- Hicks, J.B., Strathern, J.N., Klan, A.J.S. and Dellaporta, S.L. (1982) In Setlow, J.K. and Hollaender, A. (eds), *Genetic Engineering. Principles and Methods*. Vol. 4. Plenum, New York, pp. 219–284.
- Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) *J. Bacteriol.*, **153**, 163–168.
- Lin, R.-J., Newman, A.J., Cheng, S.-C. and Abelson, J. (1985) *J. Biol. Chem.*, **260**, 14780–14792.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
- Messing, J. (1983) *Methods Enzymol.*, **101**, 20–78.
- Osborne, B.I. and Guarente, L. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4097–4101.
- Platt, T. (1986) *Annu. Rev. Biochem.*, **55**, 339–372.
- Proudfoot, N.J. and Whitelaw, E. (1988) In Glover, D.M. and Hames, B.D. (eds), *Transcription and Splicing. Frontiers in Molecular Biology*. IRL Press, Oxford, pp. 97–129.
- Ruohola, H., Baker, S.M., Parker, R. and Platt, T. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5041–5045.
- Sanger, F. (1981) *Science*, **214**, 1205–1210.
- St John, T.P. and Davis, R.W. (1979) *Cell*, **16**, 443–452.
- Tajima, M., Nogi, Y. and Fukasawa, T. (1985) *Yeast*, **1**, 67–77.
- Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
- Weaver, R.F. and Weissmann, C. (1979) *Nucleic Acids Res.*, **7**, 1175–1193.
- Yu, K. and Elder, R.T. (1989) *Mol. Cell. Biol.*, **9**, 2431–2444.
- Zaret, K.S. and Sherman, F. (1982) *Cell*, **28**, 563–573.
- Zaret, K.S. and Sherman, F. (1984) *J. Mol. Biol.*, **176**, 107–135.

Received July 9, 1990; revised August 13, 1990