The yeast *FBP1* poly(A) signal functions in both orientations and overlaps with a gene promoter

Agustín Aranda¹,²,+, José Enrique Pérez-Ortín¹, Claire Moore² and Marcel.lí del Olmo¹,*

¹Departament de Bioquímica i Biologia Molecular, Facultats de Ciències, Universitat de València, and Departamento de Biotecnología, Instituto de Agroquímica y Tecnología de Alimentos, CSIC, Apartado 73, E-46100 Burjassot, Spain and ²Department of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA

Received July 8, 1998; Revised and Accepted September 7, 1998

ABSTRACT

This report provides an analysis of a region of chromosome XII in which the *FBP1* and *YLR376c* genes transcribe in the same direction. Our investigation indicates that the *Saccharomyces cerevisiae* *FBP1* gene contains strong signals for polyadenylation and transcription termination in both orientations in *vivo*. A (TA)₄₄ element plays a major role in directing polyadenylation in both orientations. While this region has four nonoverlapping copies of a TATATA hexanucleotide, which is a very potent polyadenylation efficiency element in yeast, it alone is not sufficient for full activation in the reverse orientation of a cluster of downstream poly(A) sites, and an additional upstream sequence is required. The putative RNA hairpin formed from the (TA)₄₄ element is not involved in 3′-end formation. Surprisingly, deletion of the entire (TA)₄₄ stretch affects transcription termination in the reverse orientation, in contrast to our previous results with the forward orientation, indicating that the transcription termination element operating in the reverse orientation has very different sequence requirements. Promoter elements for the *YLR376c* gene overlap with the signal for *FBP1* 3′-end formation. To our knowledge, this is the first time that overlapping of both types of regulatory signals has been found in two adjacent yeast genes.

INTRODUCTION

The initiation and the termination of mRNA synthesis are both accurately regulated during the process of transcription. The ternary complex formed on the promoter when the first nucleotide is transcribed travels through the entire gene and is disassembled downstream of the stop codon. In eukaryotes, transcription termination by RNA polymerase (pol) II is often tightly coupled to the process of cleavage and polyadenylation of the precursor RNA (1). Many studies have investigated the mRNA sequences and the trans-acting factors involved in eukaryotic polyadenylation (for reviews see 2–4). The yeast *Saccharomyces cerevisiae* has been used widely in these studies due to the convenience of its biochemistry and genetics. The ability of yeast whole-cell extracts to cleave extended precursor RNAs endonucleolytically at the same sites used for poly(A) addition *in vivo* (5–8) has indicated that, overall, the 3′-end processing reaction is similar to that of mammalian cells. It has been proposed that three types of sequence elements organized in the following order work to specify polyadenylation in *S. cerevisiae*: an efficiency element(s), which is often an alternating pyrimidine/purine stretch rich in UA pairs; an A-rich poly(A) site positioning element; and the actual site of polyadenylation, which often occurs after pyrimidine residues (9,10). However, other sequences or a different arrangement of sequences have been identified for some yeast genes (11–16). This variability can be partially reconciled if different classes of polyadenylation sites exist in *S. cerevisiae* (17).

Transcription termination downstream of a poly(A) site is thought to be important in reducing non-productive transcription and in avoiding transcriptional interference at downstream promoters (18 and references therein) and chromosomal elements such as centromeres (13,19) and origins of replication (20). In higher eukaryotes, the efficiency of transcription termination correlates with the strength of the polyadenylation signal and the presence of a termination site (21–23), which, in some cases, has been shown to pause the elongating polymerase (18,24). In the yeast *S. cerevisiae*, transcription termination occurs within 100 nucleotides downstream of poly(A) sites, consistent with the short intergenic regions found in this organism (12,25). Some of the sequence signals capable of promoting mRNA 3′-end formation in yeast are also involved directly in transcription termination (25,26). Recently, a direct association between mRNA 3′-end processing and termination of transcription has been established for the *S. cerevisiae CYC1* gene (27).

The 3′ flanks of several yeast genes, such as *CYC1* (16), *ADH1*, *ARO4*, *TRP1*, *TRP4* (17,28), *ACT1*, *YPT1* (5), the Ty retroelement (29) and *FBP1* (12), are capable of directing polyadenylation in both orientations. However, this is not a general case in yeast (30), and for genes like *GAL7* (11) and *GCN4* (31), polyadenylation is...
only possible in the forward orientation. Furthermore, some sequences behave differently in vitro than they do in vivo (5,8,11). An involvement of mRNA secondary structure has been proposed to explain the ability of some 3′ flanks to function in both orientations (8,28). In fact, inverted repeats have been described in these regions in yeast genes such as URA3 (32), ADH2 (33), GAL1, GAL7, GAL10 (8) and CYC1 (34), but a requirement for such secondary structures has not been demonstrated. The ability to direct transcription termination in both orientations has only been studied systematically for CYC1 (25,26).

The organization of 3′-end formation signals is especially interesting in areas of the genome in which two open reading frames (ORFs) are separated by a short intergenic region. This situation is quite common in the yeast genome due to its extreme compactness (35). Given that there is no preference in the orientation of two contiguous yeast genes, approximately half of the yeast genes are packed in terminator–promoter combinations (35), and mechanisms must be in place to avoid transcriptional interference at the downstream promoter. In fact, it has been shown recently that transcriptional interference does indeed occur when the natural control of transcription is modified (36).

One of these situations occurs on the yeast chromosome XII between the FBP1 gene and its closest neighbor, the YLR376c gene. Our previous analysis of the FBP1 gene (12) indicated that its polyadenylation signal is complex and that at least three efficiency elements (represented by TATA-like sequences) must be present for efficient processing in the forward orientation. This gene also shows a novel organization of sequence elements. In this report, we describe the nature of the FBP1 3′-end forming element in the reverse orientation. Our analysis has again revealed a very complex pattern of polyadenylation signals, and different requirements for transcription termination, when compared with the forward orientation. We also found that promoter elements for the downstream gene overlap sequences required for FBP1 3′-end formation. However, the efficiency of FBP1 polyadenylation is not affected by whether or not this downstream promoter is active, and initiation at this promoter takes place even when FBP1 transcription occurs through it.

MATERIALS AND METHODS

Reagents

All restriction and modifying enzymes were purchased from Amersham, Boehringer Mannheim or Promega and used according to the manufacturer’s directions. Radioactive nucleotides [32P]dATP and [32P]UTP were from Amersham and DuPont-NEN Research Products. The nylon transfer membrane Hybond-N and the Rediprime kit for random primer labeling were from Amersham. Products. The nylon transfer membrane Hybond-N and the Rediprime kit for random primer labeling were from Amersham. Products.

Plasmids

The pFBP391 plasmid (37), which contains 397 bp of the 3′ flank of the S.cerevisiae FBP1 gene cloned into pUC18, was used to create 5′ and 3′ nested deletions, and the derived plasmids were named according to the length of FBP1 sequence remaining in the original pFBP391 (12).

Derivatives of the pFBF391 and pFBP144 plasmids carrying internal deletions in the FBP1 sequence (pFBP391SM, pFBP391DM, pFBP391D1, pFBP391D2, pFBP391D3, pFBP144D2 and pFBP144D3) have been described previously (12,38).

Derivatives of the plasmid pME729 (39) were used for testing the in vivo polyadenylation ability of fragments of the 3′ region of the FBP1 gene. Fragments of FBP1 3′ end were isolated from the pFBP plasmids and inserted into the vector, as described previously (12). The resulting plasmids have been named pVPA, with a number indicating the length of the sequence in the fragment introduced. The orientation of any fragment is indicated by F (forward orientation) or R (reverse orientation).

pT7T3 18U/19U vectors from Pharmacia were used to prepare the precursors for the in vitro processing assay (12).

Yeast strains and culture methods

Yeast strain SEY2101 (a, ade2-101, leu2-3, 112, ura3-52, swc2Δ9, gal2) was a gift of Dr S. Emr and was used for the analysis of in vivo polyadenylation. It was transformed with the pVPA plasmids following the protocol of Ito et al. (40).

Yeast strains CJM88 (a, ura3) and RG1-5d (a, fbp1-2::LEU2, ura3, gal7), gifts of Dr J.M. Gancedo, were used to analyze the expression of the FBP1 and YLR376c genes.

Yeast diploid strain 1097/930 (a/α leu2/+, trpl1/tp1 prb1-1122/prb1-1122 pep4-3/pep4-3 prc1-407/prc1-407 his1+/+) was used for preparing the whole-cell extracts for the in vitro processing reactions.

All yeast culture methods were as described by Sherman et al. (41).

RNA analysis

Total RNA was isolated according to the method of Zitomer and Hall (42) using glass beads to disrupt the yeast cells. For northern analysis, 10 μg of total RNA was separated on a denaturing formaldehyde gel according to Sambrook et al. (43). After transfer to Hybond-N nylon membranes, the bound RNA was hybridized, autoradiographed and scanned as described (12).

To obtain RNAs of the CJM88 and RG1-5d strains in conditions of repression and derepression of the FBP1 gene, cells were grown in YPD medium (containing 2% glucose) to 5 × 10⁷ cells/ml. Cells were then transferred to YPE medium (containing 2% ethanol) at a concentration of 1.5 × 10⁸ cells/ml and grown overnight to achieve derepression of the FBP1 gene.

The RT–PCR to map 3′ ends of mRNAs was carried out as described in Aranda et al. (12).

For the analysis of RNA secondary structures, the MFOLD program, which is from the Genetics Computer Group (44) and based on the method of Zuker and Stiegl (45), was used. An annealing temperature of 28°C was used, and no other restriction was considered.

In vitro polyadenylation assays

Yeast whole cell extracts were prepared according to Chen and Moore (46), except that the cells were lysed with glass beads. Cleavage and polyadenylation assays with these extracts were carried out as described previously (46). Polyadenylation efficiency was determined as described by Aranda et al. (12).
RESULTS

Polyadenylation in the S. cerevisiae FBP1 gene uses TA repeats as efficiency elements in the reverse orientation

We have shown previously that a 397 bp Sau3A fragment of the 3′ end of the S. cerevisiae FBP1 gene contains the polyadenylation signal and has the ability to direct efficient 3′-end formation in both orientations (12). This fragment contains the gene sequence located between positions 1054 and 1445 (with 1 as the translation start) and includes several interesting features (Fig. 1). For instance, it contains multiple TATA-like sequences which constitute the efficiency elements of the polyadenylation signal in the forward orientation (12). Some of these repeats are organized in an almost perfect (TA) 14 region between positions 1125 and 1152, which can form a DNA cruciform in vitro (37) and is predicted to form a hairpin in the mRNA (Fig. 2B). This 397 nucleotide fragment also contains the promoter and part of the coding region of the YLR376c gene, which is transcribed in the same direction as the FBP1 gene.

To determine sequences which specify polyadenylation when the 397 nucleotide FBP1 fragment is in the reverse orientation, derivatives carrying various deletions were subcloned into pME729 plasmid (39) upstream of the ADH1 polyadenylation signal. With this construct, the proportion of readthrough transcripts to the ADH1 poly(A) site gives an estimate of the efficiency of the insert in polyadenylation. This assay was used previously to analyze the same fragments in the forward orientation (34) and is underlined. The 5′ and 3′ deletion end points are indicated by white and black triangles, respectively. The arrows and arrowheads mark the poly(A) sites in the forward and reverse orientations, respectively, as determined by RT–PCR. The coding sequence of YLR376c is written in lowercase letters.

Analysis of pVPA397F, which has the 397 nucleotide fragment in the forward orientation, shows an abundant 0.4 kb band (>98% of the mRNAs), corresponding to the use of the poly(A) sites of the FBP1 gene (12). In the reverse orientation (pVPA397R), the major band (89% of the mRNAs) is a 0.7 kb transcript which corresponds to a polyadenylation site located approximately at the end of the FBP1 insert, demonstrating polyadenylation ability in both orientations. By RT–PCR, we showed that in the reverse orientation, four poly(A) sites were scattered between the (TA) 14 repeat and the end of the insert (positions 1114, 1105, 1094, 1091), and one was located in the polycloning site of the vector from which this fragment was isolated (12). Analysis of ~100 clones for both orientations (12 and this study) indicates that frequency of nucleotides at the poly(A) site was PPy83(T70)A85, consistent with the consensus sequence, PyA, described for yeast (5).

For the following discussion, it is helpful to remember that the designation of the endpoints of the deletions as 5′ or 3′ is derived from our previous analysis of the activity of this fragment in the natural, or forward, orientation (12), and that the 5′→3′ sequence of the reverse orientation transcripts, as depicted in Figure 1, comes from reading the bottom strand from right to left. For determining the role of sequences from the 3′ part of the reverse orientation insert, we used the deletions called 5′−355, 5′−328 and 5′−307. In the first (5′−355), the efficiency drops slightly from 89 to 78% (Fig. 3A and Table 1). This decrease is most likely related to removal of several of the potential polyadenylation sites (Fig. 1) rather than removal of an efficiency element. An intermediate band (I) is detected between the truncated (T) and the readthrough (R) transcripts (Fig. 3A). This band could reflect use of a cryptic poly(A) site between the FBP1 and ADH1 terminators, which is seen only when poly(A) sites at the end of the FBP1 sequence are deleted. In the second deletion (5′−328), more poly(A) sites in the reverse orientation have been eliminated. Despite this loss, the efficiency of polyadenylation did not change relative to deletion 5′−355. The band located between the truncated and readthrough transcripts is again detected.
Figure 2. Analysis of in vivo polyadenylation of pVP A plasmids carrying 397 fragment with point mutations and internal deletions. (A) Sequence of the FBP1 gene between positions 1119 and 1159. The (TA)_{14} element is in bold. Point mutations SM and DM are indicated with lowercase letters and underlined. Internal deletions (D1–D3) are also shown. (B) Analysis of the RNA secondary structure of the main hairpin predicted by the MFOLD program in the WT and mutated sequences. The structure was obtained by folding the sequence between positions 1125 and 1152. Arrows mark point mutations in SM and DM. (C) Analysis of the effect of the SM and DM mutations in the processing efficiency of the 397 fragment in reverse orientation. The conditions and symbols are described in Figure 3A. (D) Analysis of the effect of the D1, D2 and D3 deletions in the processing efficiency of the 397 and 144 fragments in reverse orientation. The conditions and symbols are described in Figure 3A.

The fragment corresponding to the 5'-307 deletion lacks 17 bases of the (TA)_{14} sequence. There is a significant decrease in the polyadenylation efficiency of this fragment in the reverse orientation, with 58% of the transcripts being truncated (Fig. 3A and Table 1). This decrease could be due to loss of the sequence itself or to a modification of the mRNA secondary structure. The fact that greater than half of the transcripts are truncated indicates that remaining 10 nucleotide TA stretch plus upstream sequence also contributes to processing.

To investigate the role of these sequences further, we analyzed several deletions from the 5' end of the reverse orientation fragment (Fig. 3B). With the first deletion tested (3'-214), there was a decrease in the polyadenylation efficiency to 75% truncated transcripts, suggesting that the deleted sequence is stimulatory, but not essential, for polyadenylation. The processing ability is not affected by the additional reduction in 3'-161. With the next two deletions, 3'-144 and 3'-94, there is a progressive drop in efficiency, to 63 and 53%, respectively. The sequence removed in 3'-144 has no resemblance to any reported yeast efficiency elements, but the deletion of 3'-94 removes an alternating pyrimidine/purine stretch and encroaches upon the TA TA TA repeats. The greatest drop in efficiency (from 53 to 18%) comes with removal of an additional 10 bases in the 3'-84 fragment, which retains just six TA pairs of the (TA)_{14} region. These results suggest that an intact (TA)_{14} element is a major contributor to polyadenylation efficiency in the reverse orientation, especially in the absence of the additive effects of weaker upstream efficiency elements.
RNA secondary structure is not a determinant for polyadenylation in the FBP1 gene

The TA-rich region located in the 3′ end of the FBP1 gene is an important efficiency element for polyadenylation in both orientations in vivo, although activity in the reverse orientation could not be detected in vitro using yeast whole-cell extract and exogenously added RNA substrate (data not shown). To further analyze essential features of this region, we made several modifications to the TA repeat in its natural context in the 397 fragment. First, point mutations were designed to abolish the ability of the (UA)14 sequence to fold into an RNA hairpin structure (Fig. 2A). To predict the structure of the RNA of this sequence in the reverse direction, the wild-type sequence forms a predicted RNA hairpin containing only the (UA)14 element. The SM mutation introduces a central unpairing which modifies the perfect symmetry of the hairpin (Fig. 2B), causing a significant reduction in the free energy (from –6.4 kcal/mol in the wild-type sequence to –2.1 kcal/mol). The double mutation (DM, Fig. 2B) produces a complete reorganization of the whole region, giving only a very short and unstable hairpin in the UA repeat region (–0.8 kcal/mol).

These structures predict that if the global RNA structure is responsible for processing, the point mutations would not affect polyadenylation efficiency in the forward orientation, but they would progressively and adversely affect it in the reverse orientation. However, there was no loss of processing efficiency with wild-type levels of truncated transcripts in the reverse orientation of the fragment (Fig. 2C) or in the forward orientation (Table 1). These results argue against a role for the RNA hairpin in processing of transcripts in the reverse orientation. They also suggest that a perfect TA repeat in both parts of the (TA)14 stretch is not necessary for 3′ efficient processing in either orientation, and that some TA pairs are redundant.

An internal deletion of the (TA)14 region gives a reduction in both polyadenylation and transcription termination efficiencies in the reverse orientation

We have shown previously (12) that reduction of the (TA)14 element to a simple TATATA hexanucleotide still gives wild-type levels of processing in the forward orientation, but complete removal of the (TA)14 sequence greatly decreased the polyadenylation efficiency. To determine the effect of reduction of the TA content of this element on 3′-end formation in the reverse orientation, we tested the internal deletions shown in Figure 2A. The D1 fragment carries a deletion of 10 bp inside the (TA)14 region, which reduced the TA pairs from 13 to nine, the D2 fragment conserved just three TA pairs and D3 carries a complete deletion of the TA-rich sequence. When we analyzed these deletions in the reverse orientation, we found the results to be similar to those obtained for the forward orientation. Reduction of the TA content did not significantly affect the processing efficiency (89% WT, 86% D1 and 80% D2, Fig. 2D), in agreement with the idea that some of the TA pairs are redundant. With the D3 fragment, the level of truncated species dropped to 30%, confirming a major role for this sequence in the reverse direction. In this orientation, the decrease in truncated species is accompanied by an increase in the readthrough transcripts. The ratio of the total amount of RNA transcribed from the ACT1 promoter in the pVPA plasmids to that from the chromosomal ACT1 gene is similar for all four plasmids in this experiment (3.5 for the WT fragment, 3.3 for D1, 3.2 for D2 and 2.5 for D3), indicating that termination has also been affected. These results contrast with what we found previously for the D3 deletion in the forward orientation, which was defective for polyadenylation but competent for termination (12). Thus, in the reverse orientation, transcription termination as well as

Table 1. In vivo polyadenylation efficiency, in forward and reverse orientations, of the fragments analyzed in this study and in previous work (12)

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>397</td>
<td>&gt;98</td>
<td>89</td>
</tr>
<tr>
<td>5′-355</td>
<td>&gt;98</td>
<td>78</td>
</tr>
<tr>
<td>5′-343</td>
<td>86</td>
<td>n.d.</td>
</tr>
<tr>
<td>5′-328</td>
<td>89</td>
<td>79</td>
</tr>
<tr>
<td>5′-307</td>
<td>28</td>
<td>58</td>
</tr>
<tr>
<td>3′-214</td>
<td>n.d.</td>
<td>75</td>
</tr>
<tr>
<td>3′-161</td>
<td>&gt;98</td>
<td>72</td>
</tr>
<tr>
<td>3′-144</td>
<td>&gt;98</td>
<td>63</td>
</tr>
<tr>
<td>3′-122</td>
<td>85</td>
<td>n.d.</td>
</tr>
<tr>
<td>3′-114</td>
<td>53</td>
<td>n.d.</td>
</tr>
<tr>
<td>3′-94</td>
<td>47</td>
<td>53</td>
</tr>
<tr>
<td>3′-84</td>
<td>33</td>
<td>18</td>
</tr>
<tr>
<td>3′-40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>397SM</td>
<td>&gt;98</td>
<td>89</td>
</tr>
<tr>
<td>397DM</td>
<td>&gt;98</td>
<td>92</td>
</tr>
<tr>
<td>397D1</td>
<td>&gt;98</td>
<td>86</td>
</tr>
<tr>
<td>397D2</td>
<td>&gt;98</td>
<td>80</td>
</tr>
<tr>
<td>397D3</td>
<td>60a</td>
<td>30</td>
</tr>
<tr>
<td>144D2</td>
<td>&gt;98</td>
<td>36</td>
</tr>
<tr>
<td>144D3</td>
<td>72a</td>
<td>12</td>
</tr>
</tbody>
</table>

aThe values shown correspond to the standard method of calculation of the polyadenylation efficiency. In a previous study (12) we have proposed a different method of calculation that takes into account the fact that in these cases, the levels of readthrough transcripts do not increase. With this method, the efficiency is significantly lower. n.d., not determined.
polyadenylation has been disrupted by removal of the TATA
target sequence.

When the D2 deletion was made in the 3′-144 fragment, reduction
of the TA repeat had a more dramatic effect than the same deletion
in the longer 397 fragment, reducing the polyadenylation efficiency
from 65 to 36% (Table 1 and Fig. 2D). This confirms the presence
of additional activating sequences in the region between 1198 and
1445 of fragment 397, which can compensate for the partial
disruption of the TA repeat in the D2 deletion. With the D3
deletion, the efficiency dropped further to 12% and, as in the case
of the 397 fragment, there is a corresponding increase in the
readthrough transcripts.

Overlap exists between the YLR376 promoter and the
FBP1 terminator

Run-on experiments, carried out previously, have indicated that
transcription occurs through the whole 397 fragment in the
direction of the FBP1 gene but not in the opposite direction (12).
The region between the coding sequences of the FBP1 and
YLR376c genes contains just 191 bp, and the distance between the
FBP1 poly(A) site and the YLR376c start codon is only 50
nucleotides, suggesting that elements important for 3′-end
formation and transcription initiation might overlap.

To study the expression of the FBP1 and YLR376c genes, we
first analyzed mRNAs derived from the chromosomal copies of
these two genes in the strain CJM88. Because the expression of the
FBP1 gene is subjected to catabolite repression, we prepared
RNA samples from this strain grown in either repressing or
derepressing conditions. RNAs were analyzed by northern blots,
using as a probe a 2 kb fragment corresponding to the FBP1 and
YLR376c region (Fig. 4A). As expected, an FBP1 mRNA of 1.3 kb
is detected only after derepression. Another transcript of 0.95 kb
is detected under both culture conditions, and is the size predicted
for the YLR376c mRNA. It is worth noting that the part of the
filter corresponding to the YLR376c transcript is overexposed in
this figure. In fact, according to our calculations, the levels of
YLR376c transcript are about 10 times lower, relative to those of
the FBP1 transcript, under derepressed conditions of this gene.
Neither transcript is observed in the RG1-5d strain, which lacks
most of both ORFs (Fig. 4A).

To analyze the YLR376c promoter, we used some of the pVPA
plasmids employed previously for the polyadenylation competition
assays. RNAs from strains carrying these plasmids were analyzed
by northern blots, using a probe complementary to the YLR376c
ORF sequence contained in the insert. This probe will detect a
transcript (Y, 0.6 kb) initiating from the FBP1 promoter, as
well as the readthrough transcript (R, 1 kb) originating from the
ACT1 promoter (Fig. 4B). The ACT1 promoter gives a level of
transcription into the insert that is very similar to that of the FBP1
promoter in derepressed conditions (47). In spite of a high level
of transcription into the FBP1 3′ region (Fig. 3A), the YLR376c
transcript derived from the pVP397 plasmid was easily
detectable (Fig. 4B, lane 1). The level of readthrough transcript
is low but distinguishable. Deletion of 42 bp from the 5′ end of
the insert (the pVP355 plasmid) causes a complete disappearance
of the YLR376c transcript, indicating that an element required for
initiation at the YLR376c promoter has been affected. In this
construct, the polyadenylation signal of the FBP1 gene has not been
damaged (12), as can be inferred from the small amount of
readthrough transcript and wild-type levels of truncated transcript
(Table 1). Successive deletions increase the amount of readthrough
because they affect the efficiency of FBP1 polyadenylation signal
(12) and, obviously, do not permit YLR376c transcription. In the
D2 internal deletion (Fig. 4B, lane 5) the YLR376c expression is
not affected, but it drops significantly in the D3 deletion (lane 6),
suggesting that the TA repeat is also important for function of the
YLR376c promoter. Assuming that poly(A) tails in mature
In our previous study, we found that the 3′ end of the FBP1 gene is unusual because of the novel organization of its polyadenylation signal elements and the uncoupling between 3′-end formation and transcription termination (12). We also showed that polyadenylation and transcription termination are efficient in forward and in reverse orientations. In this report, we analyze the poly(A) signal in reverse orientation and show that sequences necessary for transcription of the neighboring gene are found upstream of the FBP1 poly(A) site.

The existence of polyadenylation signals working in both orientations is not a general case in yeast (30). The situation found in the CYC1 (34) and ADH1 (17) genes could be due to overlap with the signal of another gene transcribed convergently. In the case of the FBP1, the adjacent gene (YLR376c) transcribes in the same direction. While this suggests no physiological necessity for a polyadenylation site in the reverse direction, the fragment in this orientation does function as a strong polyadenylation signal in vivo (although not in vitro), and it even has a higher efficiency than the 3′ flank of the TRP1, TRP4 and ARO4 genes tested in a similar system (17). Thus, analysis of such a signal can provide further insight into the requirements for yeast polyadenylation.

Table 1 shows a comparison between the efficiency of polyadenylation in the forward (12) and reverse orientations of several fragments from the 3′ end of the FBP1 gene. By deletion analysis, we found that at least five regions were important for optimal efficiency of the flipped FBP1 3′ flank. One is the cluster of cryptic poly(A) sites activated by the presence of upstream elements. However, the cellular machinery is not very demanding about poly(A) sites, since processing of transcripts with this cluster deleted is still very high. A similar retention of efficient polyadenylation, with utilization of new sites, was found in the forward FBP1 orientation (12) and has also been reported for the CYC1 (10) and GAL7 (11) genes, when natural poly(A) sites were deleted.

Four other regions had additive effects on polyadenylation efficiency. The first two were far upstream (almost 200 nucleotides) of the poly(A) site cluster. The furthest contained a match to a TAG...TAGT motif important for the function of many other yeast poly(A) sites. The second, a 17 bp region removed in deletion 3′-144, does not correspond to any reported motifs. Another region involved in polyadenylation efficiency is identified when additional sequence is removed to create 3′-94 fragment. This deletion has eliminated three TA pairs of an almost perfect (TA)_{14} repeat located 11 bases upstream of the first poly(A), and also a TACATAA, which could function as an efficiency element (10). Finally, as might be expected, more severe encroachment into the (TA)_{14} stretch from either end (as in constructs 5′-307 and 3′-84) further decreased 3′-end processing. These deletions also demonstrated that, in this context, a single TA tract of 10–12 bases, despite its reputation as a potent signal, was not directing efficient polyadenylation. Precise excision of just the (TA)_{14} stretch caused a dramatic drop in polyadenylation efficiency, and this defect was most pronounced when the two far upstream regions were missing. In summary, the (TA)_{14} repeat contributed the most to polyadenylation in the reverse orientation, but additional upstream regions were essential for full function. This result is different from the one for the TRP4 gene. For this, 3′-end formation in the reverse orientation depended only on one TAGT motif (28). These results support the emerging model that many yeast polyadenylation signals are complex mosaics of efficiency elements of varying strength whose sequence specifications we do not fully understand. Even less is known about how the polyadenylation machinery recognizes such diverse and plentiful elements in choosing a poly(A) site.
An interesting question is whether the intrinsic symmetry of a TA repeat signal in a natural poly(A) site would always lead to function in both orientations, as happens in genes like MRPI (48) or FBPI (12 and this study). If this is the case, it is possible that the putative secondary structure adopted by these sequences is involved in the process of 3′-end formation, as has been suggested for other genes in several organisms (8,15,28,49–51). The RNA structure of the HIV-1 enhancer is a critical determinant of viral RNA synthesis in mammalian cells (52), and an RNA hairpin is involved in processing of the histone messengers in higher eukaryotes (49). However, in the case of the reverse orientation of the FBPI gene, point mutations in the (TA)14 region which create a much shorter and less stable RNA hairpin did not affect the processing efficiency. With an internal deletion that conserves only the last three TA pairs, there was no loss of efficiency in the forward direction, and only a slight decrease in the reverse orientation. Thus, a TA hexanucleotide which is not able to form a stable RNA hairpin is sufficient if other close sequences are not removed. Further analysis is required, but it seems likely that secondary structure does not play much of a role in in vivo yeast polyadenylation, and the processing factors are probably recognizing primary sequence in the mRNA rather than a structural motif.

Another interesting aspect of mRNA 3′-end formation in yeast is the relationship between polyadenylation and termination. In higher eukaryotes, the efficiency of transcription termination by RNA pol II depends on the presence of functional polyadenylation signals as well as on the strength of a transcription pause site downstream of the poly(A) site (1). Recent findings of the association of polyadenylation factors with RNA pol II transcriptional complexes (53) have lent credence to a model in which binding of these factors to nascent polyadenylation signals makes the polymerase susceptible to termination at the downstream pause site. In S.cerevisiae, a coupling between mRNA maturation and termination of transcription has been established for the CYC1 gene (27) and deletion of the sequence in this gene required for cleavage and polyadenylation has also been shown to abolish transcriptional termination in vivo (25,26). However, some mutations in the ADH2 gene affect the polyadenylation signal but not in vitro termination (54). Removal of the (TA)14 stretch in the FBPI gene did not affect transcriptional termination in forward orientation, indicating that in this example, the elements required for transcription termination are separable or do not exactly overlap with those required for polyadenylation (12). In contrast, as described in this report, deletions in the FBPI 3′ flank, which were detrimental to polyadenylation in the reverse orientation, always resulted in a corresponding increase in the level of readthrough transcripts, even for a complete removal of the (TA)14 stretch. This observation indicates that transcription termination was also adversely affected by these mutations, and leads to the conclusion that the elements which specify transcription termination are not the same in the forward and reverse orientations of the FBPI sequence. These results may not seem as puzzling if the following interpretation is considered. As we proposed previously (12), partial processing complexes may assemble on forward orientation FBPI precursor lacking the TA repeats and be sufficient to promote transcription termination but not polyadenylation. The FBPI sequence in the reverse orientation is not a natural polyadenylation signal, i.e. one optimized through evolution, and in addition may not have a good pause site associated with it. As we have shown here, the repeats of the TATATA consensus on this precursor are the major determinants of polyadenylation efficiency and, by induction, polyadenylation complex assembly. In this case, their removal would be expected to affect both processes.

A possible function for transcription termination signals downstream of poly(A) sites is to prevent RNA polymerase from entering the neighboring gene. In this respect, our results indicate that in a short intergenic region of just 191 bp, the elements specifying 3′-end formation in the FBPI gene and those needed for initiation of transcription of the YLR376c gene overlap. We found that one promoter element required for YLR376c transcription is located upstream of the positioning and efficiency elements for FBPI polyadenylation, and a second is found in the (TA)14 stretch. The location of a transcription terminator element within a yeast promoter has been shown in the case of the URA3 promoter (32), although in this case this terminator has no physiological role. In other systems, like Ty transposons in yeast (29) or mammalian virus (55,56), overlapping of terminators and promoters has also been found. The extreme compactness of the yeast genome (35) suggests that overlapping between gene ends may be a common event for this organism. A statistical analysis of intergenic distances shows that ORFs with divergent promoters are, on average, 618 bp while convergent terminators are 326 bp apart (35). This has been interpreted as 309 and 163 bp being the average size for promoters and terminators, respectively (35,36). Although these sizes fit well with experimentally determined lengths for both (25,57), they are only averages of a wide range of different sizes. A close inspection of chromosome III, for instance, reveals that many ORFs are much closer than those average values but some others are much farther. Thus, in some ORFs intergenic distances represent true non-genic regions, but for many other ORFs, the short intergenic distances mean that either the sizes of promoter and terminator are shorter than those values or they overlap. We favor the second possibility because it is difficult to conceive that short distances, such as 191 bp in the case of FBPI, could accommodate separately two promoters, two terminators or a promoter–terminator combination. The case of ARO4 and HIS7, analyzed by Springer et al. (36), in which 417 bp separate both genes, could be an example of spatially separated, non-overlapping elements. However, the case of FBP1/YLR376c, in which promoter and terminator clearly overlap is probably more frequent. A statistical analysis of the intergenic distances in chromosome III shows that 58 of the 94 promoter–terminator combinations (62%) are shorter than 472 bp (163 + 309). This raises the question of how the two transcriptional machineries cope with these kinds of situations. In natural situations, it appears that high levels of transcriptional interference have been avoided. However, low levels of interference are still possible and even tolerated. We have observed that in at least one case of overlapping terminators, that of YIL161w/POT1, there is only 56 bp between the stop codons and a small amount of interference can be detected (S.Puig, J.E.Pérez-Ortízn and E.Matallana, in preparation). Whether this is also the case for promoter–terminator combinations, such as that of FBP1/YLR376c, will require further study.

**ACKNOWLEDGEMENTS**

We thank Dr M. Kessler for his suggestions and for his support with the *in vitro* experiments, and Dr E. Matallana for critical reading of the manuscript. We also thank Drs G. Braus, J. M. Gancedo and S. Emr for the gift of plasmids and yeast
strains. We thank Servei de Bioinformàtica de la Universitat de València for technical support. This work was funded by PB91-0329 grant from DGICYT (Spanish Ministry for Education and Science) to J.E.P.-O., 1160 grant from the Universitat de València to M.d.O. and 5R01 GM4 1752 grant from the National Institute of Health of USA to C.M.

REFERENCES