Nucleic Acids Research, Vol. 20, No. 18 4733–4739

# The yeast actin intron contains a cryptic promoter that can be switched on by preventing transcriptional interference

Stefan Irniger, Christoph M.Egli, Markus Kuenzler and Gerhard H.Braus*
Institute of Microbiology, Swiss Federal Institute of Technology (ETH), Schmelzbergstrasse 7, CH-8092 Zürich, Switzerland

Received July 13, 1992; Revised and Accepted August 24, 1992

## ABSTRACT

**We show that the single intron of the actin gene of the yeast *Saccharomyces cerevisiae* contains a cryptic promoter for transcription of the second exon. This promoter is inactive in the normal actin gene, but can be activated when the actin gene promoter is deleted. An identical activation was induced by placing efficient transcriptional terminators at position 61 of the 309 bp intron. In all cases transcripts with identical 5′ ends close to the boundary of the intron and the second exon were produced. These results indicate that the cryptic promoter in the actin intron is occluded in the normal actin gene by transcriptional interference with the actin gene promoter. Transcription initiation near the intron/exon 2 boundary is enabled by protection from traversing polymerases, that initiated transcription at the upstream located actin gene promoter. A partial promoter protection using leaky terminators resulted in small amounts of transcripts initiated from the cryptic promoter. Although we do not know any function of the cryptic promoter in actin gene expression, it is tentative to speculate that the cryptic intron promoter might be a relict of a promoter that was functional earlier in evolution.**

## INTRODUCTION

The model for transcriptional interference hypothesizes that initiation of transcription is inhibited at a promoter when RNA polymerases that initiated upstream transcribe through this promoter (1, 2, 3). Traversing polymerases were suggested to reduce the activity of such a promoter and this process was therefore termed promoter occlusion. This phenomenon was initially observed in prokaryotes where it has been shown that the $P_L$ promoter of prophage λ inhibits downstream located *gal* promoters (1).

Some examples of promoter occlusion are known in eukaryotic genes, especially for promoters of ribosomal genes. A number of these RNA polymerase I promoters contain transcriptional terminators just upstream of the promoter and these terminators were shown to augment transcription initiation from the adjacent promoter (4, 5, 6, 7). Recent studies revealed that the function of the terminators is to protect the promoter from RNA polymerase I molecules transcribing through the promoter (3, 7). In the absence of a terminator the traversing polymerases impaired the attachment of new initiation complexes by damaging the complex of promoter-binding initiation factors and DNA.

Only a few RNA polymerase II promoters were shown to be affected by transcriptional interference, although many eukaryotic genes are arranged in tandem on the chromosomes and transcriptional interference was suggested to be a possible way of polymerase II gene regulation (2). An example for polymerase II promoter occlusion was identified in the avian leukosis retrovirus (ALV; 8). The promoter in the 3′ long terminal repeat (LTR) of the virus was shown to be inefficient when a functional promoter was present in the 5′ LTR. Deletion of the 5′ LTR sequences, however, activated transcription from the 3′ LTR promoter. Another experiment that demonstrated the effects of transcriptional interference of RNA polymerase II transcription was performed with an artificial tandem construct of two α-globin genes (2). In this case expression of the downstream gene was inhibited 5- to 20-fold by transcription of the upstream gene, but proper expression was restored after placing a transcriptional terminator between the two genes.

Despite these well documented observations the frequency and the significance of promoter occlusion in polymerase II gene expression remain still unclear.

In this paper we present an additional example of promoter occlusion. We show that the single intron of the *Saccharomyces cerevisiae* actin gene contains a cryptic promoter for transcription of the second exon. This promoter is inactive in the normal actin gene. If transcription across the actin intron is prevented either by insertion of transcriptional terminators or by deletion of the actin gene promoter, then initiation of trancription close to the boundary of the intron and the second exon is stimulated. Inefficient, leaky terminators that allowed partial readthrough transcription across the intron caused a partial promoter occlusion. Although no function of this cryptic promoter is known in the normal actin gene, we will discuss the possibility whether it might be a molecular fossil of a promoter that had a significance prior to the arrangement of the intron–exon structure of the actin gene. The cryptic promoter might have been a functional promoter for the independent transcription of the second actin

* To whom correspondence should be addressed

exon in early evolution, according to the theory that exons were initially independent minigenes and shuffled together later in evolution (9).

## MATERIALS AND METHODS

### Yeast strains, medias and methods

The *S.cerevisiae* strains used for these studies were RH1242 (*MAT*a *aro7 leu2-2*) and RH 1631 (*MAT*a *ura3-52*), both are derivatives of the laboratory strains X2180-1A (*MAT*a *gal2 SUC2 mal CUP1*) and X2180-1B (*MAT*α *gal2 SUC2 mal CUP1*). Yeasts were grown either in YEPD complete medium or MV minimal medium (10). Yeast transformation (11), DNA isolation (12) and Southern analysis (13) were previously described.

### Plasmid constructions

The constructions of the tests plasmid pME621 that contains the actin-*URA3* fusion gene and of the pME621-derived plasmids that contain the DNA fragments from the 3′ ends of the *GCN4*, *PHO5*, *ADH1*, *TRP1*, *TRP4* and *ARO4* genes in the intron of the fusion gene, respectively, are described elsewhere (14). The construction of a control construct containing a DNA fragment from the *TRP1* coding region is also described therein. All plasmids are 2μm derived *E.coli*/*S.cerevisiae* shuttle vectors that carry the *LEU2* gene for selection in *S.cerevisiae*.

The three plasmids which lacked the actin promoter were constructed as follows: The plasmids pME621, and its derivatives containing the *GCN4* and *TRP1* 3′ ends, respectively, were digested with *Mlu*I. This unique restriction site is located 230 bp located upstream of the transcription initiation site in the actin promoter. The linearized plasmids were subsequently incubated with nuclease Bal31 (Boehringer−Mannheim) over various time periods, blunt-ended and religated. Plasmids containing the appropriate deletions were sequenced to verify the boundaries of the deletions. The endpoints of the deletions are 365 bp (for pME621), 385 bp (for the test plasmid containing the *GCN4* 3′ end) or 260 bp (for the test plasmid containing the *TRP1* 3′ end) downstream of the *Mlu*I restriction site, respectively. Therefore in all three deletion mutants the complete actin promoter was deleted, including the transcription initiation site, but the intron sequences downstream of the *Xho*I restriction site were not affected.

The two plasmid containing the actin/*lacZ* fusions (see Figure 4A) were constructed as follows. A 343 bp *Xho*I-*Acc*I (blunt-ended) actin fragment that contains actin intron and exon 2 sequences as well as an 505 bp *Ava*II-*Acc*I (blunt-ended) fragment that contains the *GCN4* 3′ end region in addition to the actin intron/exon 2 sequences were cloned into the *Sal*I-*Hin*dIII (blunt-ended) or the *Ava*II-*Hin*dIII (blunt-ended) restriction site of the plasmid pNM480, respectively (15). This plasmid contains a promoter-less *lacZ* gene. The fusion maintained the open reading frame between actin and *lacZ* sequences.

### Isolation of RNA and Northern blotting

Total yeast RNA was isolated as previously described (16), using 0.5 mm glass beads to disrupt the yeast cells. For Northern hybridizations, 10 μg of total RNA was separated on 2% agarose gels containing 3.2% formaldehyde, transfered to nylon membranes, bound to the membrane by UV crosslinking and hybridized at 42°C with randomly labelled DNA fragments (17).

### Primer elongation

5 pmol of oligonucleotide primer was 5′-endlabeled with $\gamma$-$^{32}$P-ATP and polynucleotide kinase (Pharmacia-LKB). $10^6$ cpm of labeled primer was hybridized at 52°C to 20 μg of yeast total RNA and the primers were subsequently elongated using AMV reverse transcriptase (BioRad), according to Kassavetis and Geiduschek (18).

## RESULTS

### Deletion of the actin promoter or insertion of polyadenylation sites into the actin intron results in the production of a novel RNA

We have previously constructed the test plasmid pME621 (14), a 2μm derived plasmid that contains a gene fusion of the actin to the *URA3* gene (Figure 1A) . The actin part comprises the efficient actin promoter, the short (10 bp) first exon, the intron (309 bp) and a portion of the second exon. This partial actin gene was fused to a *URA3* cassette that contains the *URA3* gene lacking the first six amino acids (19). This test plasmid was initially used to analyze and compare several yeast polyadenylation sites (14). For this purpose DNA fragments containing the polyadenylation sites were inserted into the *Xho*I restriction site located at position +61 of the intron of the fusion gene. When analyzing the transcripts of these constructs, we observed that the polyadenylation sites caused, besides the expected truncated RNAs, the formation of an additional 1550 nucleotide (nt) transcript that was apparently located downstream of the polyadenylation sites.

We have investigated the origin of this initially unknown transcript using the plasmid constructs illustrated in Figure 1A. All constructs are derivatives of pME621. These constructs either miss the entire actin promoter (pME621-ΔP) or have an efficient polyadenylation site in the intron (G4, P5, A1), have an inefficient polyadenylation site in the intron (T1, T4, A4), have a control DNA fragment in the intron (CR) or have a polyadenylation site in the intron and miss the actin promoter (G4-ΔP, T1-ΔP).

The transcripts that were produced when these plasmids were transformed into yeast are indicated below the corresponding construct in Figure 1A and are shown in the Northern hybridization in Figure 1B. In the case of plasmid pME621 the readthrough transcripts from the actin promoter to the end of the *URA3* gene were formed. The 2000 nt unspliced pre-mRNA was efficiently spliced to a 1700 nt mRNA. When the actin promoter is completely deleted (construct pME621-ΔP) actually no *URA3* specific mRNA should be detected. Northern hybridization, however, showed that a mRNA of approximately 1550 nt is produced that hybridized to the *URA3* DNA probe.

The same transcript of 1550 nt was also detected when six different 3′ end fragments, all 200−300 bp long, containing the various polyadenylation sites, were inserted into the *Xho*I restriction site in the intron of the actin-*URA3* fusion gene. This 1550 nt transcript corresponds to our previously identified RNA of unknown origin (14). Interestingly, the 1550 nt RNA was as abundant as for construct pME621-ΔP when efficient polyadenylation sites were inserted in the intron, namely in the cases of the test plasmid containing the 3′ end regions originating from the *GCN4*, *PHO5* and *ADH1* genes (constructs G4, P5 and A1, respectively). Since the 1700 nt and 2200 nt transcripts could not be detected, these polyadenylation sites obviously prevented any significant readtrough transcription to the *URA3* gene, but

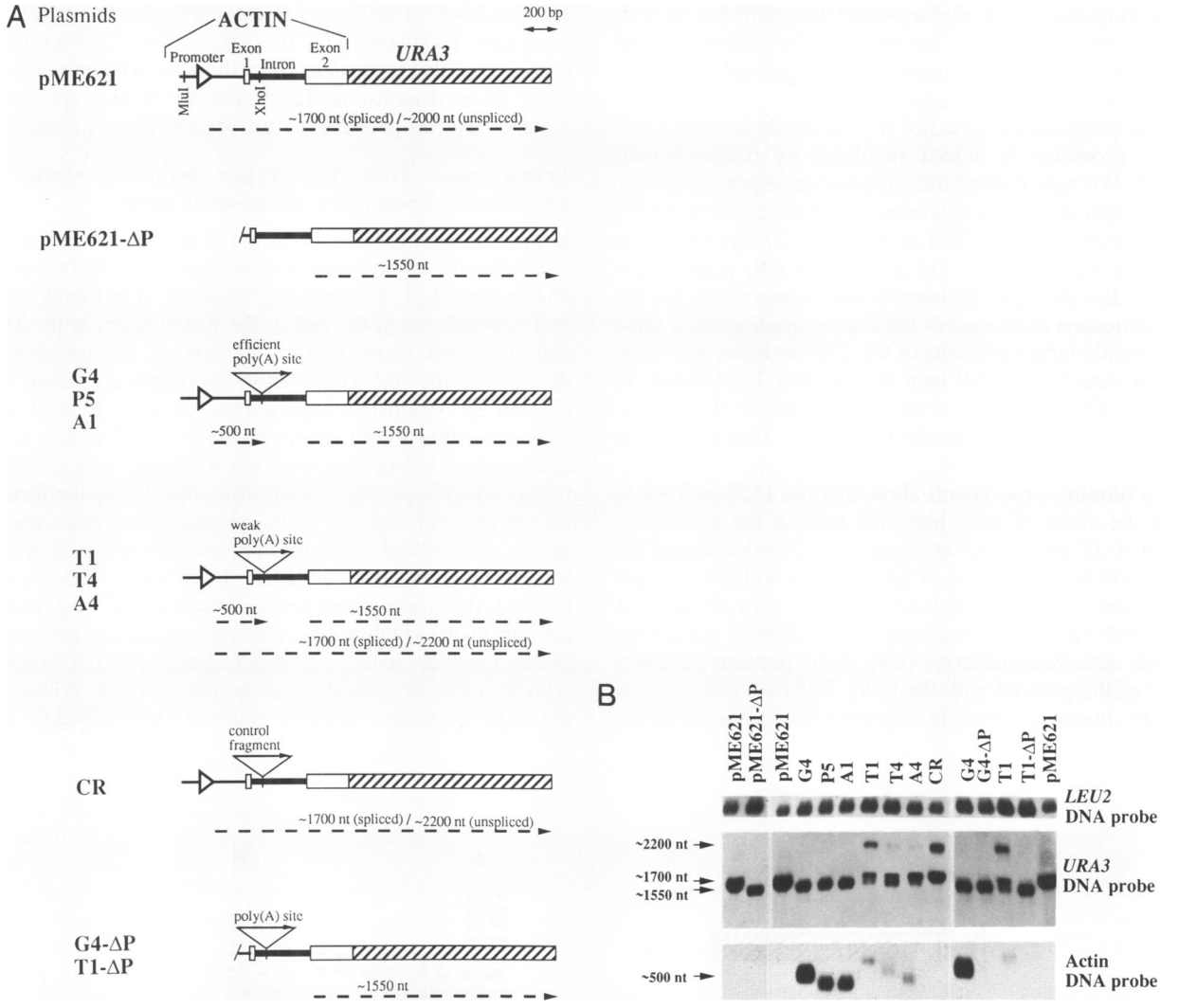


**Figure 1**. (A) The different actin-*URA3* fusion constructs. The actin-*URA3* fusions and the transcripts (indicated with dashed arrows) produced by the corresponding construct are illustrated. Plasmid pME621 (14) contains a fusion of part of the actin gene, including promoter, first exon, intron and a portion of the second exon, to a *URA3* cassette (19). The construct pME621-ΔP is similar to this plasmid except that the actin promoter was deleted (see Materials and Methods). The constructs termed G4, P5 and A1 contain three different efficient polyadenylation sites (14), originating from the genes *GCN4*, *PHO5* and *ADH1*, respectively, inserted into the *Xho*I restriction site of the intron of the fusion gene of pME621. In the constructs T1, T4 and A4 three different weak polyadenylation sites, derived from the genes *TRP1*, *TRP4* and *ARO4*, respectively, were cloned at the same location and in construct CR a control fragment originating from the coding region of the *TRP1* gene was inserted into the intron. All DNA fragments were 200−300 bp in length. The constructs G4-ΔP and T1-ΔP are identical to G4 and T1, respectively, except that the actin promoter was deleted. (B) Northern analysis of the transcripts of the various actin-*URA3* fusion constructs. 10 μg total RNA isolated from the yeast strains carrying the various plasmid constructs, designated as in Figure 1A, was separated on formaldehyde containing 2% agarose gels, transfered to nylon membranes, hybridized with a 1.1 kb *URA3 Hin*dIII-*Hin*dIII DNA probe and exposed. After removal of the radioactivity, the membrane was hybridized to a 0.45 kb actin *Mlu*I-*Xho*I DNA probe and then to a *LEU2* DNA probe in order to standardize the amount of the RNA loaded. The larger mRNAs (2200 and 1700 nt) hybridizing to the actin probe are not shown. The transcripts produced are indicated with arrows and correspond to the transcripts drawn in (A).

resulted in the formation of more or less equal amounts of truncated, actin-specific RNAs and of the 1550 nt transcript. In contrast, the 1550 nt RNA was present in only small amounts in the cases of the test plasmids containing the less efficient *TRP1*, *TRP4* and *ARO4* polyadenylation sites (constructs T1, T4 and A4, respectively) that still allow partial readthrough transcription (2200 and 1700 nt mRNAs). A construct containing a DNA fragment derived from the coding region of the *TRP1* gene that did not function as a poly(A) site (construct CR), produced only the readthrough transcripts, but neither the truncated actin nor the 1550 nt RNA. These results demonstrate that the appearance of the 1550 nt RNA was dependent on the presence of a functional polyadenylation site in the intron or on the absence of the actin gene promoter. The amount of the 1550 nt RNA was corresponding to the efficiency of the polyadenylation sites (Figure 1B).

## Formation of the 1550 nt RNA is regulated by transcriptional interference

The Northern hybridization experiments with the promoter deletion mutant and the constructs containing the polyadenylation sites indicate that the 1550 nt transcript is due to a cryptic promoter that is located downstream of the *Xho*I cloning site in the actin intron. Both, deletion of the actin gene promoter and

insertion of polyadenylation sites activated this promoter. In both cases this activation might be due to the prevention of transcriptional interference. Inefficient polyadenylation sites resulted in a weak promoter activation.

Additional constructs were tested in order to demonstrate that the cryptic promoter is indeed regulated by transcriptional interference. We have deleted the complete actin gene promoter region in the test plasmids containing either the efficient *GCN4* (construct G4) or the leaky *TRP1* (construct T1) polyadenylation sites, creating the constructs G4-ΔP and T1-ΔP, respectively (Figure 1A). If transcriptional interference is responsible for the only weak activation in the case of the leaky polyadenylation sites, then significantly larger amounts of the 1550 nt RNA should be produced in construct T1-ΔP than in construct T1, because in construct T1-ΔP lacking the actin gene promoter the cryptic promoter is completely protected from transcriptional interference.

Northern blotting experiments show that the 1550 nt RNA is present in the cases of both plasmids lacking the actin gene promoter (G4-ΔP and T1-ΔP in Figure 1B). The deletion of the actin gene promoter in the case of the construct with the efficient *GCN4* polyadenylation site (G4-ΔP) does not alter the amount of the 1550 nt transcript comparing to the plasmid containing the complete actin gene promoter (G4). Actin promoter deletion in the case of the plasmid with the leaky *TRP1* polyadenylation site (T1-ΔP), however, strongly augments the formation of the 1550 nt RNA as compared to the same plasmid with an active actin gene promoter (T1). Instead no more readthrough RNA is produced, due to the deletion of the actin gene promoter. These observations demonstrate that the activity of the cryptic promoter is dependent on protection from transcriptional interference.

### Transcription from the cryptic promoter starts at the intron/exon 2 boundary of the actin gene

In order to characterize the origin of the 1550 nt transcript and to localize the cryptic promoter, we have identified the endpoints of this transcript. We found by nuclease S1 mapping that its 3′ end is positioned at the end of the fusion gene, at the *URA3* 3′ end (data not shown). The 5′ ends of the transcript were determined for each construct using primer extension analysis (Figure 2). Extension experiments were performed with an oligonucleotide primer encompassing the DNA region where the actin and the *URA3* gene were fused (Figure 2A), therefore the primer hybridized only to transcripts encoded by the fusion gene on the plasmid, but not to the chromosomal actin and *URA3* genes. Analysis of the extension products received with yeast total RNA revealed for pME621 the expected 410 nt extension product, that corresponds to the transcriptional start point in the actin promoter (Figure 2B). The same signal was present for the plasmid with the inactive control fragment (CR). The constructs with the efficient polyadenylation sites (G4, P5, A1) as well as the actin promoter deletion mutants (pME621-ΔP, G4-ΔP,

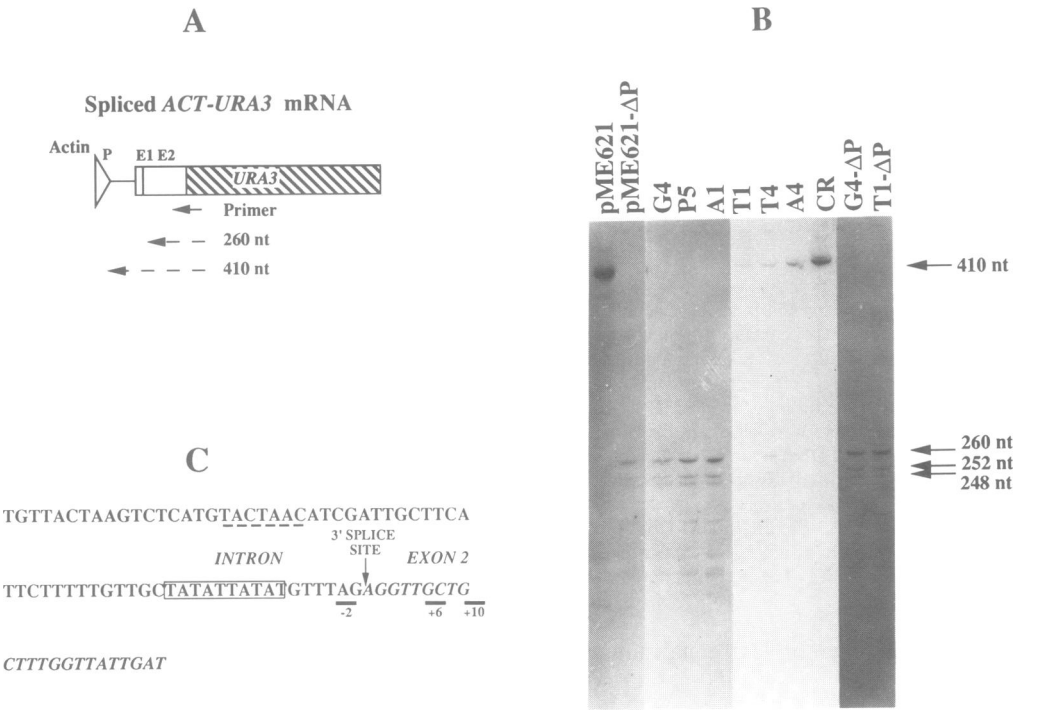


**Figure 2.** Mapping of the mRNA 5′ ends of the various actin-*URA3* fusion constructs. (A) Illustration of the spliced actin-*URA3* mRNA, the oligonucleotide primer and of the major elongation products. The primer spanned the region where the actin and the *URA3* gene were fused, its sequence is 5′-GAAAGGTCCAAAGCTTGGATCT-3′. The 3′ ends of the 248, 252 and 260 nt elongation products are located at the beginning of exon 2, the 3′ end of the 410 nt product is located in the actin promoter. Note that the intron including the cloned DNA fragments therein are removed during splicing. (B) Primer extensions experiments. $10^6$ cpm (0.1 pmol) of the 5′ endlabeled primer was hybridized to 20 μg total RNA from the yeast strains carrying the various plasmid constructs. The constructs are designated as explained in Figure 1A. The arrows indicate the major extension products. (C) Sequence upstream and downstream of the intron / exon 2 boundary (20). Exon 2 sequences are typed in italics. Significant sequence elements in this region are emphasized: underlined with dashed line is the sequence TACTAAC in the intron that is essential for lariat formation in the splicing process and indicated with an arrow is the boundary of the intron and the second exon, the 3′ splice site. A possible sequence for recognition of the TATA binding protein TFIID is boxed. The nucleotides where the major 5′ ends of the 'downstream' RNA were mapped are underlined (corresponding to the 260, 252 and 248 nt extension products in (B)). The numbers indicated the distance from the intron/exon 2 boundary.

T1-ΔP) caused smaller extension products of 248−260 nt as expected from the Northern blotting experiments. These signals correspond to the 5′ end of the 1550 nt RNA and are positioned close to the boundary of the intron and the second exon of the actin gene, at positions −2, +6, +10 with respect to the beginning of the second exon (Figure 2C). In the cases of the inefficient polyadenylation sites (constructs T1, T4 and A4) extension products ending either at the intron/exon 2 boundary or in the actin promoter were detected (Figure 2B), thereby supporting the Northern blotting results showing the readthrough as well as the 1550 nt RNA. For all constructs that induced the formation of the 1550 nt transcript, the 5′ ends were at identical positions, close to the intron/exon 2 boundary, 240 bp downstream of the *Xho*I site, where the polyadenylation sites were cloned. The promoter sequences responsible for transcription initiation at the intron/exon boundary must therefore be situated upstream of the initiation site, obviously in the actin intron. Figure 2C shows a possible TATA box (21), immediately upstream of the initiation site.

### Actin intron sequences induce the formation of a *lacZ* specific mRNA

We have further investigated the promoter function of the actin intron sequences using a promoter test plasmid, that contains a promoter-less *lacZ* gene (15). Insertion of intron sequences in front of the *lacZ* gene is expected to result in *lacZ* specific mRNA, in the case that the intron sequences indeed act as a promoter. In addition, β-galactosidase activity should be measured if the gene fusion is properly translated.

A DNA fragment spanning the actin intron promoter region and part of the second exon was fused in frame to the *lacZ* gene (Figure 3A) and this gene fusion was integrated into the yeast chromosome. A second, similar fusion construct, but that contains in addition the *GCN4 poly(A)* site upstream of the actin intron sequence was also integrated. RNA from yeast strains that harboured these constructs on the chromosome was analyzed in Northern blotting experiments (Figure 3B). Both constructs induced the formation of *lacZ* specific mRNAs. Multiple *lacZ* transcripts were produced. This might be due to multiple *lacZ* mRNA 3′ ends located in different regions of *lacZ*, an observation previously reported by another group (22). Alternatively, the 2.6 kb and the smaller signals might be *lacZ* mRNAs tied up in ribosomal RNAs. These results support our theory that the intron sequence acts as a transcription initiation sequence. However, no distinct β-galactosidase activity of the fusion gene could be measured, indicating that the actin-*lacZ* transcript is translated very inefficiently, although an ATG initiation codon for translation of the actin-*lacZ* fusion gene is present in the actin sequence. The reasons for the deficiency in translation are unknown.

The *GCN4* polyadenylation site upstream of the intron does not significantly augment the amount of mRNA indicating that the terminator upstream of the transcription initiation site is only required for proper promoter function, if transcription from an upstream promoter would otherwise inhibit it.

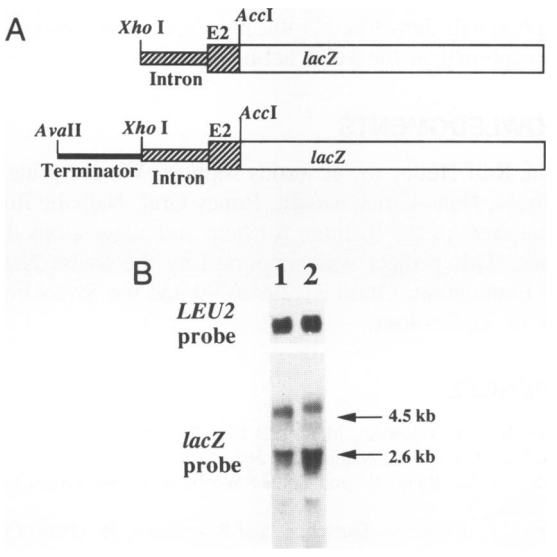


**Figure 3.** The actin intron sequences tested for promoter function in a *lacZ*-fusion. (A) Two different DNA fragments were fused to the *lacZ* gene (15) carrying at its 3′ end the yeast *ADH1* polyadenylation site, 1.) a *Xho*I-*Acc*I actin fragment containing 248 bp of the intron and 98 bp of the second exon and 2.) a similar *Ava*II-*Acc*I fragment containing in addition the 260 bp *GCN4* polyadenylation site. (B) These fusions were integrated at the *URA3* locus into the yeast chromosome and transcripts were analyzed in Northern blotting experiments using a *lacZ* DNA probe. *LacZ* specific transcripts are indicated by arrows. The 4.5 kb RNA corresponds to the complete transcript of the fusion gene. The filter was subsequently hybridized with a *LEU2* probe as a standard for the loaded RNA.

## DISCUSSION

The actin gene is one of the few genes of the yeast *Saccharomyces cerevisiae* that contains an intron. This single intron is located at the extreme 5′ end of the gene, in a way that the first exon is only 10 bp in length. In this paper we have shown that sequences within this intron are able to initiate transcription very close to the beginning of the second exon. We have shown that this 1550 nt RNA is the product of a cryptic promoter located in the actin intron.

These cryptic promoter sequences are only functional when protected from RNA polymerases traversing it. In the normal

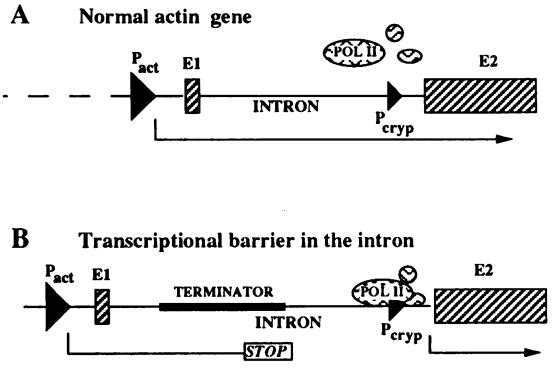


**Figure 4.** Model system for occlusion of the cryptic promoter in the actin intron. (A) In the normal actin gene, transcription initiates at the actin promoter ($P_{act}$) and transcribes unaffected the first exon (E1), the intron and the second exon (E2). Initiation at the cryptic promoter ($P_{cryp}$) at the end of the intron is inhibited by passing polymerases. (B) The introduction of efficient terminators in the intron prevent that RNA polymerases pass through the intron. This allows the formation of initiation complexes of RNA polymerase II and associated factors at the cryptic promoter in the intron and transcription initiates close to the beginning of the second exon.

actin gene the promoter in the intron seems to be inactive, as seen for plasmid pME 621 (Figures 1B and 2B) and as in previously reported transcript analysis of the actin gene (23). RNA polymerase II molecules pass at a high frequency unaffected across the actin gene, due to the efficiency of the actin gene promoter.

Activation of the cryptic promoter could be achieved by two distinct ways of protecting it from traversing RNA polymerases, namely by either removing the entire actin gene promoter or by placing polyadenylation sites between the actin gene promoter and the cryptic promoter.

If the polyadenylation sites prevent transcription of the sequences downstream of them, then they actually also have to act as transcriptional terminators. This theory has been proposed for the yeast *CYC1* polyadenylation site that was shown to cause transcription termination within a short distance (< 100 bp) downstream of it (24, 25). The *GCN4*, *PHO5* and *ADH1* fragments obviously are efficient polyadenylation sites and terminators, the *TRP1*, *TRP4* and *ARO4* fragments would be inefficient polyadenylation sites and terminators (14). While the efficient terminators completely prevent transcriptional interference, the inefficient terminators act only partially, as can be seen in the weak signals for the 1550 nt product (Figure 1B). We could not exactly determine the efficiency of the weak terminators, but recent investigation of the *TRP4* terminator revealed that only about 10% of transcription from the actin gene promoter reads through this terminator (N.Rüf, S.Irniger and G.H.Braus, unpublished results).

A model for activation of transcription at the actin intron promoter is illustrated in Figure 4. A transcriptional barrier upstream of the intron promoter prevents read-in of upstream initiated polymerase molecules into the intron promoter, and in this case the initiation complex consisting of RNA polymerase II and associated transcription factors (for a review see 21) can bind unaffected to the promoter sequences. Traversing polymerases prevent or disrupt the formation of these initiation complexes. The promoter is then occluded.

The activation and inhibition of the transcription initiation site in the actin intron is one of only few examples for transcriptional interference for RNA polymerase II promoters described up to now. Our findings with the promoter in the actin intron support previously postulated models for the effects of transcriptional interference on promoter activity (2, 8).

We do, however, not know any function of this promoter in actin gene expression. In a previous report a low level of RNA was detected that also initiates within the yeast actin intron, but that is transcribed in the antisense direction of the actin gene, towards the actin promoter (26). No meaning could be assigned to this transcript.

We can not rule out the possibility that the sequences enabling initiation of transcription are present in the intron just by chance. Alternatively, the cryptic promoter might be a relict of a promoter that was functional earlier in evolution, prior to the arrangement of the intron–exon structure of the actin gene.

The debate about the evolutionary origin and function of the introns is still contradictory. According to one theory introns were present since the genome originated and were important in piecing together the genomes of the primordial organisms, an evolutionary process often referred to as exon shuffling (9, 27, 28), while according to a second theory introns arose and were integrated into the genome relatively late in evolution (29, 30).

The identification of a promoter in the actin intron and of transcription initiation sites close to the beginning of the second exon might indicate that the second exon of the actin gene has been transcribed early in evolution independently from the first exon. When considering the proposal that exons were initially independent minigenes (9), the exons of the actin gene would initially have been spread in the genome. The cryptic promoter would then have been the initiation site for a transcript of the second exon. Later in evolution the different exons might have been joined to a single transcription unit by exon shuffling and the promoter of the second exon, that was still present in the intervening sequence between the two exons, was no longer required and was inactivated by transcriptional interference from the promoter located upstream of the first exon. A recent report about an intron in a clam provided evidence that some introns may have evolved from 5′ non-coding regions of a gene (31). In a similar manner the actin intron might be derived from a 5′ non-coding region of an ancestor gene.

The opposing theory that the actin intron was gained late in evolution was recently supported by a comparison of the actin intron positions in 23 different species (32). The introns of this highly conserved gene are randomly distributed in the various organisms and therefore were suggested to have arosen late in evolution. Introns might have originated from mobile DNA that integrated into functional genes. For example in maize, several transposable elements are spliced like introns from pre-mRNA (33). The intron in the yeast actin gene might be a relict of a transposable element and this transposable element could have introduced the cryptic promoter. Retroelements like retroviruses or the yeast Ty retrotransposon contain promoters on both ends of their genome (34).

Thus the origin of the cryptic promoter in the actin intron might serve as an indicator for the origin of the actin intron. Testing of other intron sequences for their ability to initiate mRNA transcription will show whether this phenomenon occurs generally or is exceptional in the yeast actin intron.

## ACKNOWLEDGMENTS

We thank Ralf Hütter for generous support. We are grateful to David Jones, Hans-Ulrich Mösch, Roney Graf, Nathalie Rüf and other members of the institute for help and suggestions during this work. This project was supported by the Swiss National Science Foundation, Grant 31.29926.90 and the Swiss Federal Institute of Technology.

## REFERENCES

1. Adhya, S. and Gottesman, M. (1982) Cell 29, 939–944.
2. Proudfoot, N.J. (1986) Nature 322, 562–565.
3. Henderson, S., Ryan, K. and Sollner-Webb, B. (1989) Genes Dev. 3, 212–223.
4. Grummt, I., Kuhn, A., Bartsch, I. and Rosenbauer, H. (1986) Cell 47, 901–911.
5. Henderson, S. and Sollner-Webb, B. (1986) Cell 47, 891–900.
6. McStay, B. and Reeder, R.H. (1986) Cell 47, 913–920.
7. Bateman, E. and Paule, M.R. (1988) Cell 54, 985–992.
8. Cullen, B.R., Lomedico, P.T. and Ju, G. (1984) Nature 307, 241–245.
9. Gilbert, W. (1987) Cold Spring Harbor Symp. Quant. Biol. 52, 901–905.
10. Miozzari,G., Niederberger, P. and Hütter, R. (1978) J.Bacteriol. 134, 48–59.
11. Ito, H., Jukuda, Y., Murata, K. and Kimura, A. (1983) J. Bacteriol. 153, 163–168.
12. Braus, G., Furter, R., Prantl, F., Niederberger, P. and Hütter, R. (1985) Arch. Microbiol. 142, 383–388.

13. Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
14. Irniger, S., Egli, C.M. and Braus, G.H. (1991) Mol. Cell. Biol. 11, 3060–3069.
15. Minton, N.P. (1984) Gene 31, 269–273.
16. Furter, R., Paravicini, G., Aebi, M., Braus, G., Prantl, F., Niederberger, P. and Hütter, R. (1986) Nucleic Acids Res. 14, 6357–6373.
17. Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266–267.
18. Kassavetis, G.A. and Geiduschek, E.P. (1982) EMBO J. 1, 107–114.
19. Alani, E. and Kleckner, N. (1987) Genetics 117, 5–12.
20. Ng, R. and Abelson, J. (1980) Proc. Natl. Acad. Sci. USA 77, 3912–3916.
21. Roeder, R.G. (1991) Trends Biochem. Sci. 16, 402–408.
22. Barnes, C.A., Johnston, G.C. and Singer, R.A. (1991) Gene 104, 47–54.
23. Gallwitz, D. and Seidel, R. (1980) Nucleic Acids Res. 8, 1043–1059.
24. Osborne, B.I. and Guarente, L. (1988) Genes Dev. 2, 766–772.
25. Russo, P. and Sherman, F. (1989) Proc. Natl. Acad. Sci. USA 86, 8348–8352.
26. Thompson-Jäger, S. and Domdey, H. (1990) Curr. Genet. 17, 269–273.
27. Darnell, J.E. and Doolittle, W.F. (1986) Proc. Natl. Acad. Sci. USA 83, 1271–1275.
28. Doolittle, W.F. (1978) Nature 271, 581.
29. Rogers, J.H. (1989) Trends Genet. 5, 213–216.
30. Cavalier-Smith, T. (1991) Trends Genet. 7, 145–148.
31. Naito, Y., Riggs, C.K., Vandergon, T.L. and Riggs, A.F. (1991) Proc. Natl. Acad. Sci. USA 88, 6672–6676.
32. Dibb, N.H. and Newman, A.J. (1989) EMBO J. 8, 2015–2921.
33. Wessler, S.R. (1989) Gene 82, 127–133.
34. Elder, R.T., Loh, E.Y. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 2432–2436.