Nucleic Acids Research, 1997, Vol. 25, No. 2 417–422

Sequence requirements of the bidirectional yeast *TRP4* mRNA 3′-end formation signal

Christoph M. Egli, Katrin Düvel, Nathalie Trabesinger-Rüf+, Stefan Irniger and Gerhard H. Braus*

Institute of Microbiology, Georg-August University, Grisebachstraße 8, D-37077 Göttingen, Germany

Received July 15, 1996; Revised and Accepted November 13, 1996

ABSTRACT

The yeast *TRP4* 3′-end formation signal functions in both orientations in an *in vivo* test system. We show here that the *TRP4* 3′-end formation element consists of two functionally different sequence regions. One region of ~70 nucleotides is located in the untranslated region between the translational stop codon and the major poly(A) site. The major poly(A) site is not part of this region and can be deleted without a decrease in *TRP4* 3′-end formation. 5′ and 3′ deletions and point mutations within this region affected 3′-end formation similarly in both orientations. In the center of this region the motif TAGT is located on the antisense strand. Point mutations within this motif resulted in a drastic reduce of 3′-end formation activity in both orientations. A second region consists of the 3′-end of the *TRP4* open reading frame and is required for 3′-end formation in forward orientation. A single point mutation in a TAGT motif of the *TRP4* open reading frame abolished *TRP4* mRNA 3′-end formation in forward orientation and had no effect on the reverse orientation.

INTRODUCTION

Yeast messenger RNA 3′-end formation resembles the process in mammals in many aspects. It has been shown for several yeast mRNAs that mRNA 3′-end formation requires cleavage of a precursor RNA and polyadenylation. There are, however, some remarkable differences between mRNA 3′-end formation in mammals and in yeast. Whereas cleavage and polyadenylation are strongly coupled in mammals (1), they can be uncoupled in yeast (2). No universal sequence motif exists that corresponds to the AATAAA polyadenylation signal in higher eukaryotes but various different sequence motifs involved in mRNA 3′-end formation have been identified in the yeast *Saccharomyces cerevisiae*. These motifs TAGTA, TATATA, TACATA, TATGTA, TACGTA (3–5), TTTATA, TATGTT, TATTTA (6), TTTTTATA (7) and $(AT)_9$ (8,9) show a high content in A and T nucleotides. In initial studies by Zaret and Sherman (5) a tripartite sequence motif TAT...TA(T)GT...TTT has been proposed to be involved in 3′-end formation. In the cauliflower mosaic virus (CaMV) a condensed version of the tripartite motif has been shown to be essential and we have shown that an exchange from a T to a C nucleotide within such motifs reduces 3′-end formation activity (10). In different studies it has been shown that several of these sequence elements act in concert to result in full activity (6,10,11). Signal sequences for mRNA 3′-end formation are degenerate and present in several copies at the 3′-ends of RNA polymerase II transcribed genes in yeast. In contrast to the highly conserved mammalian hexanucleotide sequence motif where single point mutations mostly abolish function, only one single point mutation was reported to dramatically reduce 3′-end formation in an authentic gene of the yeast *S.cerevisiae*. This point mutation was found in the *ADH2* 3′-untranslated region (3′-UTR) (12).

We have shown that 3′-end formation elements in yeast can be grouped into two functionally different classes (13). The 3′-end formation elements of one class direct 3′ processing only in one orientation (unidirectional) in an *in vivo* test system, whereas 3′-end formation elements of the other class act in both orientations (bidirectional). A representative of the class of unidirectional 3′-end formation elements is the *GCN4* site. This element is highly complex and contains two copies of the signal sequence TTTTTAT interrupted by a region which is responsible for poly(A) site selection (14,10). The 3′-end regions of genes from the class of bidirectional 3′-end formation sites, *TRP1*, *TRP4* and *ARO4*, all contain at least one TA(T)GT sequence motif on their sense as well as on their antisense DNA strand. We have proposed that these motifs might be a prerequisite for the bidirectional function of these elements (13). This prompted us to analyse the effects of mutations within these elements.

In this report, we have analyzed and defined the bidirectional *TRP4* 3′-end formation element in both orientations in more detail. We were interested how mutations within this region would affect 3′-end formation function in both orientations. We hypothesized that mutations resulting in different effects on polyadenylation for each orientation would suggest the existence of two distinct 3′-end formation signals for each orientation. Similar effects of mutations on both orientations support the existence of a single 3′-end formation signal.

We found that the *TRP4* 3′-end formation element can be subdivided into two different regions. The essential sequence motifs of both regions are TAGT sequence motifs. One region is located between the translational stop codon and the major poly(A) site. The poly(A) site itself is not required for function. Most changes within this region affect mRNA 3′-end formation in both orientations. Changes in the other region affect mRNA

*To whom correspondence should be addressed. Tel: +49 551 393770; Fax: +49 551 393793; Email: gbraus@gwdg.de

+Present address: Institute of Organic Chemistry, Swiss Federal Institute of Technology (ETH), Universitätsstrasse 16, CH-8092 ZYrich, Switzerland

3′-end formation only in the forward orientation. Surprisingly, this region is located at the end of the *TRP4* open reading frame (ORF) which is a novel feature of yeast 3′-end formation sites.

MATERIALS AND METHODS

Yeast strains, media and methods

The yeast strain RH1376 has been previously described (10) and is a derivative of the *S.cerevisiae* laboratory standard strain X2180-1A (*MAT***a** *gal2 SUC2 mal CUP1*). Yeast strains were cultivated in YEPD complete medium or MV minimal medium (15). Yeast transformation (16), DNA isolation (17) and Southern analysis (18) were previously described.

Enzymes and oligonucleotides

Restriction enzymes were purchased from Boehringer (Mannheim, Germany) and New England BioLabs (Schwalbach, Germany). Vent DNA polymerase was purchased from New England BioLabs. Oligonucleotides were synthesized by Microsynth (Balgach, Switzerland).

Plasmid construction and cloning

A 270 base pair (bp) fragment from the *Bst*EII to the *Eco*RV of the *TRP4* 3′-end region was amplified using the polymerase chain reaction (19) and the two oligonucleotide primers NAT1 and NAT2 containing *Xho*I restriction sites at their ends. This fragment was cleaved with *Xho*I and inserted in either orientation into the *Xho*I digested plasmid pME729 (10). The resulting vector was the basis for the construction of the various *TRP4* point mutations in this work. All point mutations, small deletions and insertions were created using oligonucleotide primers that contained the desired mutations by the polymerase chain reaction according to Giebel and Spritz (20). The 5′ and 3′ deletions were constructed as follows. The *TRP4* element was amplified from position –111 to +191 using appropriate oligonucleotides, restricted with *Xho*I and *Eco*RV, and cloned into the *Xho*I/*Sma*I sites of the vector pGEM7®-7Zf(+) (Promega, Madison, WI, USA). The plasmid was linearized with *Bam*HI for the 5′ deletions or with *Xba*I for the 3′-end deletions and treated with the *Bal*31 exonuclease enzyme. The ends of the shortened plasmids were made blunt and ligated with a *Pst*I linker [5′-d(GCTGCA-GC)-3′] for the 5′ deletions and with a *Pst*I and *Hin*dIII linker [5′-d(CCAAGCTTGG)-3′] for the 3′ deletions. The screened deletion fragments were isolated using the restriction enzymes *Pst*I and *Hin*dIII and cloned in both orientations into the test construct on the plasmid pME729 which was cut with the enzymes *Pst*I and *Hin*dIII. All constructs were verified by DNA sequencing (21).

RNA analysis

Total RNA from *S.cerevisiae* was isolated according to Furter *et al.* (22). For Northern hybridization experiments, ~10 µg total RNA was precipitated, resuspended and denatured in 30 µl sample buffer [50% v/v deionized formamide, 6% v/v formaldehyde, 1× loading buffer, 10% (v/v) TE buffer] for 5 min at 65°C and put on ice. The RNA was separated on a denaturing formaldehyde agarose gel. The 1.4% w/v agarose gel [3% (v/v) formaldehyde, 20 mM MOPS, 5 mM Na-acetate, 1 mM EDTA) was run for 3 h at 60 V in a buffer containing 20 mM MOPS, 5 mM Na-acetate and 1 mM EDTA. The gel was soaked twice in 25 mM Na-phosphate buffer for 20 min each time and the RNA was transferred onto a nylon membrane (Amersham, UK) by electroblotting (2 A, 50 V) for 3–4 h in 25 mM Na-phosphate buffer. After washing in 2× SSC, drying on 3MM paper and crosslinking under UV light (254 nm) for 5 min, the membrane with the bound RNA was hybridized at 42°C with a labeled fragment for 24 h in 50 ml hybridization mix [50% v/v formamide, 50 mM Na-phosphate pH 6.5, 800 mM NaCl, 1 mM EDTA, 0.5% (w/v) SDS, 10× Denhardt's solution, 150 µg/ml calf thymus DNA, 500 µg/ml torula yeast RNA]. The 440 bp *Mlu*I–*Xho*I DNA fragment of the *ACT1* 5′ region was randomly radiolabeled as described (23). The RNA was visualized by autoradiography. Band intensities from autoradiographs were quantified using a phosphoimager (Molecular Dynamics, Sunnyvale, CA, USA).

RESULTS

The *TRP4* 3′-end formation element causes 3′-end formation in an orientation-independent manner in an *in vivo* test system (13). An additional puzzling characteristic of this 3′-end formation site is that no *in vitro* processing could be detected (data not shown) using various yeast cell extracts which were able to process transcripts containing other polyadenylation sites derived from *GCN4*, *CYC1* or *ADH1* (14).

This prompted us to ask (i) what mutations within this element would affect 3′-end formation function in both orientations, and (ii) what sequence requirements are responsible for this bidirectional function? The *TRP4* gene 3′-UTR contains three TA(T)GT(T/A)(T) sequences, originally proposed by Zaret and Sherman (5) to be involved in 3′-end formation. One motif is located 12 bp downstream of the translational stop codon (13). The other two motifs are located in inverse orientation 24 and 49 bp downstream of the A of the translational stop codon (Fig. 1). The most upstream motif TATGTT was named ZS1, the proximal inverse TAGTTT motif was referred to as ZS2, and the distal TAGTA motif ZS3. In addition, a TAGT motif is located in the *TRP4* ORF starting at position –17.

To analyze the *TRP4* 3′-end formation element in more detail, we constructed 5′ and 3′ deletions as well as specific point mutations (Figs 1, 3 and 4). The effects of all modifications were determined on the transcript level by performing Northern blot analysis. The modified *TRP4* 3′-end elements were cloned in both orientations into the *Xho*I restriction site of the intron of the test system (Fig. 1). This gene consists of the strong *ACT1* promoter containing the 5′ part of the intron, and the *ADH1* 3′-end formation element (10). The resulting constructs were transformed into the uracil-deficient *S.cerevisiae* strain RH1376 and total RNA was analysed in Northern hybridization experiments. Efficient 3′ processing elements inserted into the *Xho*I site promote cleavage and polyadenylation of all transcripts resulting in truncated transcripts (T). Non-functional elements result exclusively in readthrough transcripts (RT). The 3′-end formation activities were quantified by measuring the band intensities by phosphoimager analysis. The ratio between the amount of truncated and the amount of the readthrough transcripts indicates the efficiency of the 3′-end formation element. The quotient of the amount of truncated transcripts to the total amount of transcripts i.e. T/(T+RT) was calculated. The wild-type *TRP4* 3′-end

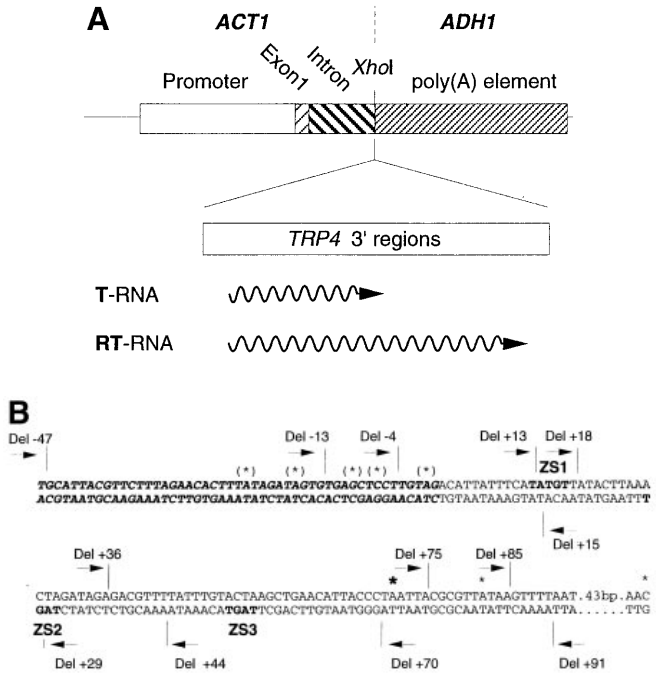

Figure 1. Test cassette for quantitative analysis of 3′-end formation elements *in vivo*. (**A**) The test gene consists of a fusion of the 5′ region of the *ACT1* gene (promoter, exon 1 and part of the intron) and the *ADH1* 3′-end formation element. Various modifications of the *TRP4* 3′-end region were inserted in either orientations into the *Xho*I site of this plasmid. The short truncated (T) and the long readthrough (RT) transcripts produced from the *ACT1* promoter are indicated with arrows. (**B**) Partial DNA sequence representing the 3′-end region of the *TRP4* gene (8). The *TRP4* ORF is written in bold italic letters, and the numbers correspond to the assignment of the A nucleotide of the TAG translational stop codon as position 0. The sequences TATGTT on the upper sense strand and the sequences TAGTA...TAGTTT on the lower antisense strand are written bold. The major and the two minor poly(A) sites for the positive orientation are indicated with a bold star and with normal stars, respectively (8,16). The poly(A) sites of the negative orientation are indicated with stars in parentheses. Vertical lines indicate the end of the various 5′ and 3′ deletions. The numbers represent the distance of the deletion endpoints to the A nucleotide of the TAG translational stop codon with the value 0. The horizontal arrows show the direction of the deletions that were performed from restriction sites upstream or downstream of the 3′ terminal *TRP4* element by digestion with the exonuclease *Bal*31. All the mutations were inserted in both orientations into the test construct.

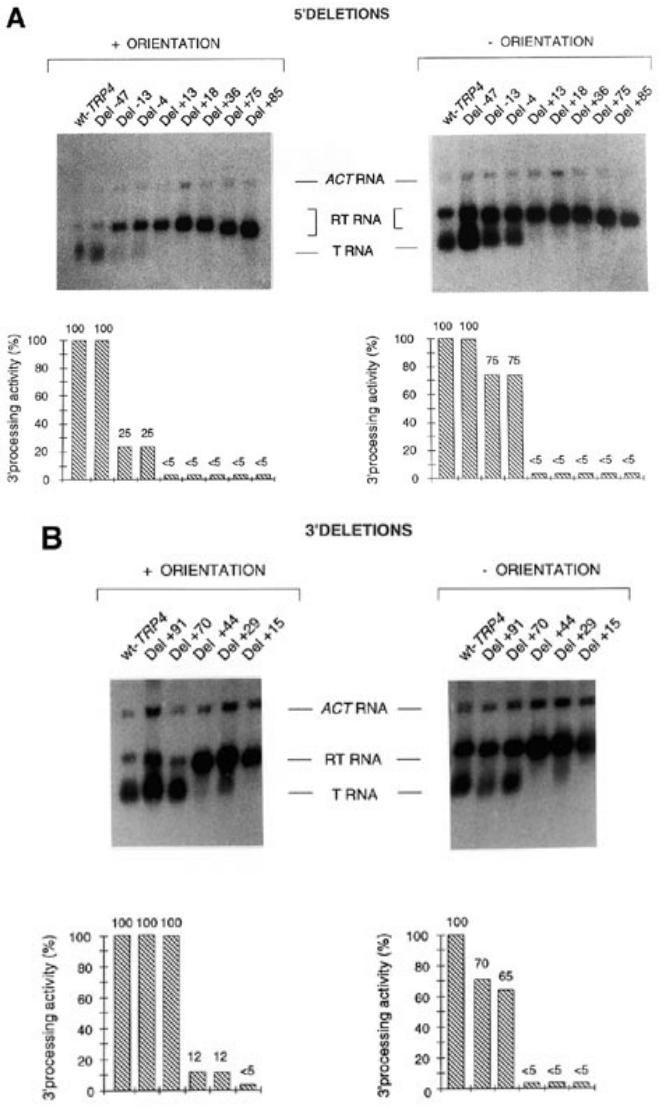

Figure 2. Analysis of the *TRP4* 3′ element by 5′ and 3′ deletion. (**A** and **B**) The effects of these deletions were determined by Northern hybridization with total RNA from yeast strains carrying the various plasmids. The designations of the mutations correspond to Figure 1B. The truncated RNAs (T) were between 300 and 380 nt, and readthrough RNAs had a length of ~720 nt. Transcripts were visualized with a radiolabeled 0.45 kb *Mlu*I–*Xho*I DNA probe from the *ACT1* promoter region. Thus the *ACT1* transcript was also visualized and used as standard. The 3′-end formation efficiencies were determined by using a phosphoimager. The numbers above the columns in the diagram represent in percent the activities of 3′-end formation which is calculated by dividing the amount of truncated transcripts by the total amount of transcripts. All values represent mean values of at least three independent Northern blots. The standard deviation did not exceed 10%.

formation activities showed different efficiencies depending on the orientation in the test system. The *TRP4* polyadenylation element was more efficient in forward than in reverse orientation. In forward orientation, 85% of all transcripts initiating from the *ACT1* promoter were truncated (T-RNA) and ended in the *TRP4* fragment. Only 15% were readthrough transcripts (RT-RNA) and ended at the *ADH1* poly(A) site of the test gene. In reverse orientation the ratio of truncated to readthrough RNA was 60%:40%. The 3′-end formation efficiencies of the wild-type *TRP4* element were used as standard and defined as 100% in either positive and negative orientation (Figs 2, 3 and 4).

The UTR of the *TRP4* mRNA 3′ element and a region of the ORF are necessary for efficient 3′-end formation in both orientations and in forward orientation, respectively

A 5′ and a 3′ deletion set of the 260 bp *TRP4* 3′-end formation element was created using exonuclease *Bal*31. Eight deletions with 5′-end points at positions –47, –13, –4, +13, +18, +36, +75 and +85 relative to the A residue of the translational stop codon at position 0 were created (5′ deletions; Fig. 1). The endpoints of five 3′ deletions were at positions +91, +70, +44, +29 and +15 (3′ deletions; Fig. 1).

Activities of 3′-end formation in forward and reverse orientation were not impaired if the *TRP4* 3′-end formation element lacked sequences upstream of position –47 (Del –47). The 5′ border of the minimal and fully functional element is therefore located in

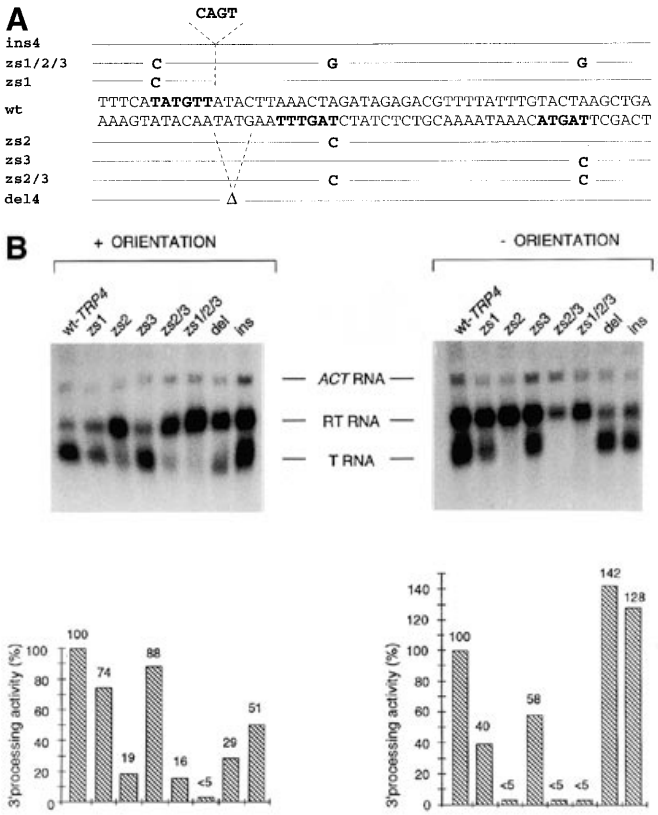

Figure 3. Influence of the TATGTT element (ZS1) of the sense strand and of the two TAGT(A/T)(T) elements (ZS2 and ZS3) of the antisense strand on 3′-end formation. (**A**) The first T nucleotide of the TATGT motif ZS 1 was mutated to a C nucleotide, resulting in the mutation zs1. Furthermore the motifs ZS2 and ZS3 were mutated by exchanging in both sequence elements the first T nucleotide to a C nucleotide (zs2 and zs3). In the mutations zs2/3 both sequence elements, and in mutation zs1/2/3 all three TA(T)GT(A/T)(T) motifs were modified simultaneously. Furthermore the spacing between the ZS1 and the ZS2 motifs was modified either by inserting the 4 nt CAGT (ins4) or by deleting the 4 nt TATG (del4). Horizontal lines represent wild-type sequences, whereas the triangle and the dotted line stand for sequence deletion. All the mutations were inserted in both orientations into the test construct. (**B**) The effect of the mutations were determined as described in Figure 2. The numbers above the columns in the diagram represent in percent the 3′ processing activity which is calculated by dividing the amount of truncated transcripts by the total amount of transcripts. All values represent mean values of at least three independent Northern blots. The standard deviation did not exceed 10%.

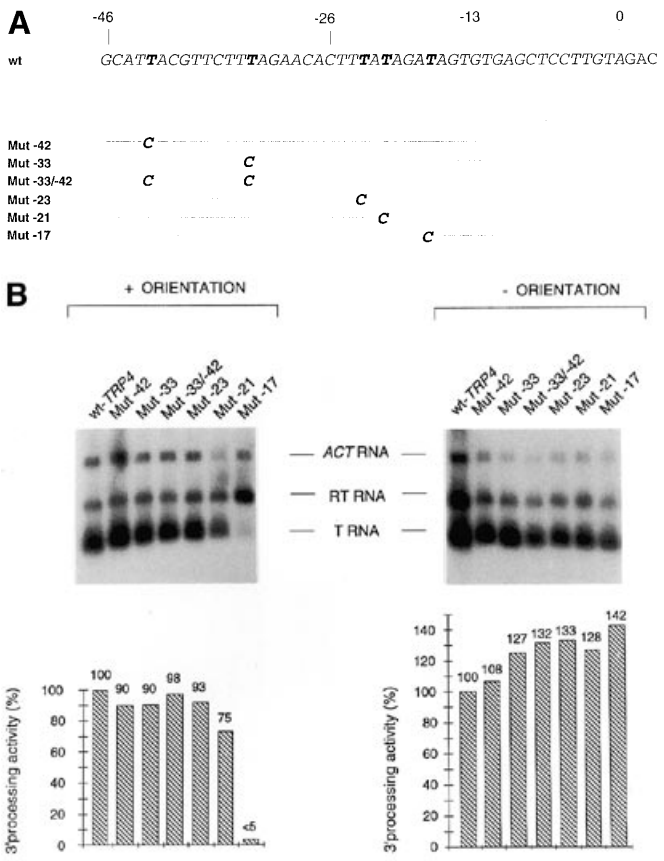

Figure 4. Analysis of the involvement of ORF elements in 3′ end formation. (**A**) The sequence element between the positions –46 and +3 is shown. Bold letters symbolize nucleotide exchange. The five T nucleotides at positions –17, –21, –23, –33 and –42 were exchanged against a C nucleotide by point mutagenesis resulting in the mutations Mut –17, Mut –21, Mut –23, Mut –33 and Mut –42. Mut –33 and Mut –42 were combined (Mut –33/–42). All mutations were inserted in both orientations into the test construct. (**B**) The effect of the mutations were determined as described in Figure 2. The numbers above the columns in the diagram represent in percent the 3′ processing activity which is calculated by dividing the amount of truncated transcripts by the total amount of transcripts. All values represent mean values of at least three independent Northern blots. The standard deviation did not exceed 10%.

the *TRP4* ORF at a position between nucleotides (nt) –47 and –13 (Fig. 2A).

The 3′-end formation activities in positive orientation in the deletions Del –13 (and Del –4) within the *TRP4* ORF dropped significantly to only one fourth of the wild-type activity. In case of the reverse orientation, however, these deletions had only minor effects on the polyadenylation activities. Therefore some sequences within the *TRP4* ORF between the positions –47 and –13 seem to be required for mRNA 3′-end formation in forward orientation. Polyadenylation activity was completely lost if sequences up to position +13 were deleted. These deletions affected the polyadenylation function in both orientations.

These results gave a first hint that the *TRP4* 3′-end formation element can be divided into regions which are required for function in both orientations and which are located in the untranslated part of the *TRP4* gene and in regions which are only required for function in forward orientation and which are located in the translated part of the *TRP4* gene. This finding was surprising since ORF sequences have not been reported to be involved in 3′-end formation in yeast so far. This part of the *TRP4* ORF region was investigated in further detail in order to determine the important sequence motifs (see below).

3′ Deletions were used to determine the 3′ border of the minimal fully active *TRP4* polyadenylation element. Activities of 3′-end formation were strongly reduced in both orientations when sequences downstream of position +44 were deleted. In positive orientation, deletion of all sequences downstream of position +70 including the major (at position +82) and two minor poly(A) sites was without effect. Thus, in forward orientation the border of the minimal fully functional element was between the positions +70 and +44. In reverse orientation efficiency of 3′-end formation was slightly reduced from 100 to 70% if the region between position +161 and +91 was deleted. Deletion of sequences between +70 and +44 led to a complete loss of activity.

In summary, our results defined the borders of a minimal significantly functional bidirectional *TRP4* 3′-end formation element between positions –47 and +70 relative to the A nucleotide of the *TRP4* translational stop codon.

The central TAGT motif in the UTR is crucial for *TRP4* 3′-end formation in both orientations

Sequence analysis revealed three sequence motifs TA(T)GT-(A/T)(T) named ZS1, ZS2 and ZS3 present in the *TRP4* 3′-UTR (Fig. 1B). Data from the 5′ and 3′ deletion experiments indicated that all three ZS motifs are located within the region which is required for polyadenylation function in both orientations. Motif ZS1 was partly destroyed in the 5′ deletion Del +13 in which activity was abolished. Motif ZS3 was deleted in the 3′ deletion Del +44 that abolished 3′-end formation (Fig. 2). The ZS2 motif is located in the center of this region (Fig. 1). To test whether these three motifs are required for 3′-end formation in both orientations, we point-mutated all three ZS motifs simultaneously in a triple mutation as indicated in Figure 3 (zs1/2/3). The first T nucleotides in the TA(T)GT(A/T)(T) sequences of all ZS motifs were exchanged against a C nucleotide since a similar T to C mutation was shown to abolish mRNA 3′-end formation of the CaMV 3′-end formation signal in yeast (10). The triple point mutation resulted in an abolishment of *TRP4* 3′ processing activity in both orientations. This result suggests an important role of the ZS motifs for *TRP4* 3′-end formation. Single point mutations in all ZS motifs were constructed to define their individual role for mRNA 3′-end formation more precisely. A change in ZS1 from TATGTT to CATGTT resulted in an activity reduction to 74% in forward and to 40% reduction in reverse orientation (Fig. 3). Mutations in both ZS motifs on the antisense strand (ZS1 and ZS2) by replacement of the first T nucleotide in the motifs against a C nucleotide resulted in drastic effects. The 3′-end formation activity decreased to 16% in forward and to loss of function in reverse orientation. A single point mutation in the central ZS2 motif revealed that this motif was the most important determinant involved in bidirectional *TRP4* 3′-end formation. Activity of this single point mutation was decreased to 19% in forward and completely abolished in reverse orientation. A single point mutation in the sequence motif ZS3 showed similar effects as a mutation in ZS1 in that 3′ processing activity was reduced to 88% in forward and to 58% in reverse orientation (Fig. 3).

All mutations in all ZS motifs resulted in significant effects in both orientations suggesting that they all play a role in the bidirectional feature of the *TRP4* polyadenylation site. We constructed a small insertion and a small deletion in the DNA sequence between ZS1 and ZS2 to test whether the distance between the ZS elements is important for 3′-end formation. We either inserted the 4 nt CAGT at position +18 (ins4) or deleted the 4 nt TATG from position +19 to +22 (Fig. 3). The 3′-end formation activity was strongly affected by these spacing modifications and was reduced in forward and enhanced in reverse orientation. In forward orientation it was reduced 3-fold to 29% in case of the deletion and 2-fold to 51% in case of the insertion. In reverse orientation activity was increased to 142% for the deletion mutation (del4) and to 128% in case of the insertion mutation (ins4). Therefore these changes affect the bidirectional function of the 3′-end formation site and support the idea that beside ZS2 as essential determinant the whole UTR is involved in 3′-end formation in both orientations.

The TAGT motif in the *TRP4* ORF is crucial for mRNA 3′-end formation in forward orientation

The 5′ deletion analysis indicated that ORF sequences located between the positions –47 and –13 are required for *TRP4* 3′-end formation in the forward orientation (Fig. 2). Since ORF sequences have not been described so far to be involved in RNA 3′ processing, we aimed to identify the determinants important for this activity. This effect does not seem to be due to changed mRNA stability. Quantification of the transcripts resulted in no change of the total amount of transcripts. The decrease of the amount of truncated RNAs (T-RNA) corresponded to an increase of the amount of readthrough RNA (RT-RNA). Therefore, the observed effects seem to be due to differences in the efficiency of mRNA 3′-end formation.

We searched for special sequences to further identify the important ORF determinants for *TRP4* 3′-end formation. In a first approach we decided to concentrate on the TAGT motif and the three sequences of this region where more than three A or T residues are located adjacent to each other. We mutated the Zaret Sherman core sequence TAGT at positions –17/–14 to **C**AGT. In addition, we mutated the sequence AT**T**A at positions –44/–41 to AT**C**A, the sequence TT**T**A at positions –35/–32 to TT**C**A, and the sequence TT**T**A**T**A at positions –25/–20 to TT**C**ATA and TTTA**C**A, respectively (Fig. 4).

With one exception (Mut –17) the point mutations did not reduce activity to a large extent in the forward orientation (between 98 and 75%). The activities were even increased (between 108 and 133%) when inserted in reverse orientation into the test gene. However, 3′-end formation activity was completely abolished if the T nucleotide of the TAGT motif at position –17 was exchanged against a C nucleotide (Mut –17). An activity loss was only observed if the element was cloned in the forward orientation. This suggests that the Zaret-Sherman sequence TAGT in the *TRP4* ORF is an important determinant of the *TRP4* polyadenylation site required for function in the forward orientation.

DISCUSSION

The sequence requirements for yeast mRNA 3′-end formation sites have been studied for a long period of time. The analysis has been difficult because many yeast elements seem to be complex consisting of numerous redundant elements. For the 3′-end formation site of the yeast *GCN4* gene encoding a transcriptional regulator we have recently shown that ~100 nt are required for fully active 3′-end formation function (14). An additional difficulty in the analysis of yeast polyadenylation sites is the variety of degenerate sequence motifs and their different functions. Several lines of evidence suggest at least three different roles for sequence motifs in 3′-end formation. One has to differentiate between efficiency elements of 3′-end formation with various TA-rich sequence motifs (3,5,11,14), elements of various A-rich sequences which select the poly(A) addition site (14,24), and the actual poly(A) site itself which seems to prefer PyA sequences (25).

In this study we have focused on the capacity of the *TRP4* element to direct 3′-end formation in both orientations which is a special feature of numerous yeast polyadenylation sites. Our aim was to find out what sequence determinants are required for a yeast 3′-end formation site to be able to function in both orientations. With our experiments we could not discriminate between the various processes in mRNA 3′-end formation

including processing, polyadenylation and termination of the precursor transcript. We wanted to know whether there is a common feature for both orientations in such 3′-end formation sites or whether we have simply two independent and overlapping sites. Bidirectional function is a characteristic for numerous but not for all yeast 3′-end formation sites. We have analyzed the *TRP4* 3′-end formation element as an example for such a symmetrical element. Other bidirectional 3′-end formation elements are the Ty retroelement (26) and the *ARO4*, *TRP1*, *CYC1*, *GAL1*, *GAL7* and *GAL10* polyadenylation elements (13,27). An example of a 3′-end formation signal of the other class is the complex *GCN4* polyadenylation site (14) which functions very efficiently but exclusively in the natural forward orientation. The significance of bidirectional signals for mRNA 3′-end formation in yeast is still unclear. No convergent transcripts have been identified downstream of the *TRP4* and of the *GAL* genes (13,27). Bidirectional 3′-end formation signals might have a safety function in the cell, namely to prevent that any RNA polymerases transcribe from the 3′-end into a gene and interfere with transcription of this gene.

Our results have shown that the bidirectional *TRP4* 3′-end formation element consists of two different parts. The region of ~70 nt located in the 3′-UTR between the translational stop codon of the *TRP4* ORF and the major poly(A) addition site is required for 3′-end formation in both orientations. The site of poly(A) addition is not necessary for efficient 3′-end formation in either orientation. In this region a TAGT(T) motif (ZS2) in inverse orientation is flanked by a proximal TATGT motif (ZS1) and a distal TAGTA motif (ZS3) (Fig. 1). The central TAGTT motif is essential for 3′-end formation in both orientations and does not tolerate point mutations without drastic loss of function in both orientations. A single point mutation destroying 3′-end formation function in forward orientation was observed up to now in only one yeast gene i.e. in the *ADH2* 3′-end formation element (12). For the *GCN4* 3′-end formation element which we have analysed recently (14) we could not find such drastic effects on 3′-end formation function. This suggests that the unidirectional *GCN4* element seems to include more redundant signals for interaction with the 3′-end formation machinery than the *TRP4* element. One possibility to explain the result that the point mutations within the ZS motifs of the *TRP4* element as well as 5′ and 3′ deletions of this region had significant effects in both orientations is the existence of a higher order RNA structure defined by the *TRP4* element. The ZS2 TAGT motif is the essential element of this region. It might be located at an exposed position and seems to be a good candidate for an interaction site between RNA and the proteins of the 3′-end formation machinery. The ZS1 and ZS3 motifs seem to contribute as auxiliary elements. Small changes in the spacing between ZS1 and ZS2 affect function in forward and reverse orientations differently. Whereas 3′-end formation function in forward orientation is significantly decreased, the function in the reverse orientation is even improved. We assume that differences in spacing destroy a higher order RNA structure of this region required for the bidirectional function.

Additional sequences located in the 3′ terminal part of the *TRP4* ORF are required for 3′-end formation in forward orientation. ORF sequences have not been reported so far to be involved in 3′-end formation *in vivo*. For the *GAL1* and *GAL10* 3′-end formation elements 3′ processing was reported to be reduced in an *in vitro* reaction when ORF sequences were deleted (27). Presumably this effect was dependent on differences in the transcript length of various transcripts tested. In the case of the *TRP4* ORF element transcript length has not been changed. The effect of a loss of polyadenylation function specifically in forward but not in reverse orientation was caused by a single point mutation. Interestingly this point mutation was also located in the ZS core motif TAGT. Exchanges in all other TA-sequences in the 3′ terminal part of the ORF did not have any significant effect on 3′-end formation.

Both TAGT motifs in the ORF as well as ZS2 in the untranslated region are essential for the function in forward orientation. In reverse orientation 3′-end formation seems to depend only on one motif, the ZS2 TAGT motif in the UTR. This might be one reason why the efficiency of mRNA 3′-end formation in the test system of the wild-type element in the reverse direction is lower when compared to the forward orientation.

ACKNOWLEDGEMENTS

We thank Christoph Springer and Chris Berens for critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft and by the Fonds der Chemischen Industrie.

REFERENCES

1 Wahle,E. and Keller,W. (1992) *Annu. Rev. Biochem.* **61**, 419–440.
2 Egli,C.M. and Braus,G.H. (1994) *J. Biol. Chem.* **269**, 27378–27383.
3 Irniger,S. and Braus,G.H. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 257–261.
4 Russo,P., Li,W.Z., Hampsey,D.M., Zaret,K.S. and Sherman,F. (1991) *EMBO J.* **10**, 563–571.
5 Zaret,K.S. and Sherman,F. (1982) *Cell* **28**, 563–573.
6 Guo,Z., Russo,P., Yun,D.F., Butler,J.S. and Sherman,F. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4211–4214.
7 Henikoff,S. and Cohen,E.M. (1984) *Mol. Cell. Biol.* **4**, 1515–1520.
8 Abe,A., Hiraoka,Y. and Fukasawa,T. (1990) *EMBO J.* **9**, 3691–3697.
9 Chen,J. and Moore,C. (1992) *Mol. Cell. Biol.* **12**, 3470–3481.
10 Irniger,S., Sanfaon,H., Egli,C.M. and Braus,G.H. (1992) *Mol. Cell. Biol.* **12**, 2322–2330.
11 Russo,P., Li,W.Z., Guo,Z. and Sherman,F. (1993) *Mol. Cell. Biol.* **12**, 7836–7849.
12 Hyman,L.E., Seiler,S.H., Whoriskey,J. and Moore,C.L. (1991) *Mol. Cell. Biol.* **11**, 2004–2012.
13 Irniger,S., Egli,C.M. and Braus,G.H. (1991) *Mol. Cell. Biol.* **11**, 3060–3069.
14 Egli,C.M., Springer,C. and Braus,G.H. (1995) *Mol. Cell. Biol.* **15**, 2466–2473.
15 Miozzari,G., Niederberger,P. and Hütter,R. (1978) *J. Bacteriol.* **134**, 48–59.
16 Ito,H., Jukuda,Y., Murata,K. and Kimura,A. (1983) *J. Bacteriol.* **153**, 163–168.
17 Braus,G., Furter,R., Prantl,F., Niederberger,P. and Hütter,R. (1985) *Arch. Microbiol.* **142**, 383–388.
18 Southern,E.M. (1975) *J. Mol. Biol.* **98**, 503–517.
19 Oste,C. (1988) *BioTechniques* **6**, 162–186.
20 Giebel,L.B. and Spritz,R.A. (1990) *Nucleic Acids Res.* **18**, 4947.
21 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
22 Furter,R., Paravicini,G., Aebi,M., Braus,G., Prantl,F., Niederberger, P. and Hütter,R. (1986) *Nucleic Acids Res.* **14**, 6357–6373.
23 Feinberg,A.P. and Vogelstein,B. (1984) *Anal. Biochem.* **137**, 266–267.
24 Guo,Z. and Sherman,F. (1995) *Mol. Cell. Biol.* **15**, 5983–5990.
25 Heidmann,S., Obermaier,B., Vogel,K. and Domdey,H. (1992) *Mol. Cell. Biol.* **14**, 4633–4642.
26 Yu,K. and Elder,R.T. (1989) *Mol. Cell. Biol.* **9**, 2431–2444.
27 Sadhale,P.P. and Platt,T. (1992) *Mol. Cell. Biol.* **12**, 4262–4270.