

Different positioning elements select poly(A) sites at the 3'-end of *GCN4* mRNA in the yeast *Saccharomyces cerevisiae*

Katrin Düvel and Gerhard H. Braus*

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

Received August 19, 1999; Revised and Accepted October 28, 1999

ABSTRACT

Cleavage and polyadenylation of eukaryotic mRNA requires efficiency and positioning elements in the 3'-untranslated region (3'-UTR) of the mRNA. Specific point mutations were introduced into the yeast *GCN4* 3'-UTR to detect sequence motifs which are involved in the positioning of the poly(A) site. 3'-End processing activities of different *GCN4* 3'-UTR alleles were measured in an *in vivo* test system. Point mutations in an AAGAA motif defocused selection of the poly(A) sites of the *GCN4* 3'-UTR to various additional poly(A) sites instead of the single site of the wild-type *GCN4* 3'-UTR. A strain with an intact wild-type *GCN4* 3'-UTR but impaired in *RNA15* encoding an RNA-binding processing factor showed a similar defocused pattern of poly(A) site selection. Remarkably, two additional sequence motifs upstream of the AAGAA motif which resemble yeast efficiency motifs independently affected poly(A) site positioning but not efficiency of 3'-end processing. Mutations in one motif resulted in an additional upstream poly(A) site. Alterations of the other motif shifted the poly(A) sites exclusively to two downstream poly(A) sites. These data suggest several contact points between the precursor mRNA and the polyadenylation machinery in yeast.

INTRODUCTION

Most eukaryotic mRNAs carry a poly(A) tail which is added to specific cleavage sites in the 3'-untranslated region (3'-UTR) in a post-transcriptional process (for recent reviews see 1,2). The sequence requirements for efficient 3'-end processing are well defined in higher eukaryotes. A highly conserved AAUAAA motif is located upstream and a GU-rich region downstream of the poly(A) site. The positioning of the poly(A) site is determined by the distance between these two sequence motifs. In yeast, less conserved sequence motifs vary in their arrangement. Recently, a computer alignment of 3'-UTRs of 861 different yeast genes was performed to enlighten the question of sequence motifs (3). A model consisting of five different signal sequences was proposed based on the frequency of

sequence motifs at specific positions in the 3'-UTRs. Three sequence motifs are located upstream of the poly(A) site which itself is one of the five elements and most often found to be a PyA dinucleotide. The far upstream efficiency motif and the positioning motif located further downstream were previously confirmed by experimental data (4,5). The computer analysis revealed an additional T-rich region in close proximity upstream of the poly(A) site and a motif localized downstream of the poly(A) site. However, experimental studies of different yeast 3'-UTRs showed that each region has its unique features, including specific sequence motifs and their peculiar arrangement in the 3'-UTR (4,6–9). The yeast *CYC1* gene contains one of the best-studied 3'-ends with independently functioning efficiency and positioning motifs. The AAGAA positioning motif of the *CYC1* 3'-UTR can be deleted without affecting the processing efficiency (10). A total mutagenesis revealed AATAAA and AAAAAA as optimal positioning motifs (11). An (A)₁₁ motif directs the positioning of poly(A) sites of the *FBP1* gene (9). In the case of the *FBP1* 3'-UTR the positioning motif is located upstream of the efficiency motif. Deletion of the positioning motif results in a more scattered pattern of poly(A) sites due to additional downstream sites (9). An entirely different positioning motif was found in the *TRP4* 3'-UTR. Its sequence composition, TAGT, resembles efficiency motifs of other yeast genes. Correspondingly, a point mutation within this motif not only altered poly(A) site selection but also drastically reduced the 3'-end processing activity (12). An additional difference to the formerly described positioning motifs was the localization of new poly(A) sites exclusively upstream of the wild-type poly(A) site due to a point mutation in the TAGT motif.

Besides the *cis* sequences, a complex protein machinery is necessary for mRNA 3'-end processing in eukaryotes. There are at least three different complexes in yeast, each consisting of several protein factors which are involved in this process. They are called PFI for polyadenylation factor I, and CFI and CFII (cleavage factors I and II). An additional factor, which might also be part of the PFI complex, the poly(A) polymerase Pap1p, has been described (reviewed in 1,13). Rna15p is part of the CFI complex and is a quite well-studied processing factor. It binds RNA with a preference for a U-rich consensus sequence deduced from *in vitro* selection experiments (SELEX) (14,15). In strains carrying a temperature-sensitive allele of *RNA15*, an alteration in poly(A) site selection could be observed at the restrictive temperature. The *ACT1* 3'-end,

*To whom correspondence should be addressed. Tel: +49 551 39 3771; Fax: +49 551 39 3820; Email: gbraus@gwdg.de

which regularly uses five different poly(A) sites, was taken as example. The temperature shift led to alterations in the frequency with which each poly(A) site was used (16).

In this study, the positioning motifs of the complex *GCN4* 3'-UTR were analysed. A former deletion analysis of the *GCN4* 3'-UTR suggested two TTTTAT sequences as efficiency motifs for 3'-end processing (7). The region in between these two sequence motifs was also necessary for full processing activity. There are several TA-rich sequence motifs located in this region which resemble efficiency motifs of other genes. In addition, a deletion of a part of this region altered selection of the poly(A) site (7). Since an additional AAGAA motif, a putative positioning element which has not been analysed so far, is also present in this region, we wondered whether the 3'-UTR of *GCN4* contains several positioning elements. Here, we show that alterations in the AAGAA sequence *in cis* resulted in similar defocussed mRNA 3'-ends as found with an intact *GCN4* 3'-UTR and defects in Rna15p function *in trans*. In addition, positioning of the *GCN4* 3'-end could also be changed in the presence of an intact AAGAA by introducing single point mutations elsewhere in the *GCN4* 3'-UTR. We found point mutations which resulted in shorter as well as longer mRNA species. Therefore, we propose the existence of several contact points between the precursor mRNA and the polyadenylation machinery.

MATERIALS AND METHODS

Yeast strains

The *in vivo* test system vectors carrying different alleles of the *GCN4* 3'-UTR were transformed into the yeast strain RH1631 (MATa, *ura3-52*) which is a derivative of the *Saccharomyces cerevisiae* laboratory standard strain X2180-1A (MATa *gal2 SUC2 mal CUP1*). The test system cassette was integrated at the *URA3* locus. The following strains were derived from RH1631: RH2258 (*ura3-52::pME1566*) carrying *GCN4mut-1*; RH2259 (*ura3-52::pME1567*) with *GCN4mut-2*; RH2260 (*ura3-52::pME1568*) with *GCN4mut-3*; RH2261 (*ura3-52::pME1569*) with *GCN4mut-4*; RH2262 (*ura3-52::pME1570*) with *GCN4mut-p*. The strains RH2395 (MATa, *ura3-1::pME800*, *trp1-1*, *ade2-1*, *leu2-3,112*, *his3-11,15*, *rna15-2*), carrying the wild-type *GCN4*, and RH2411 (MATa, *ura3-1::pME1568*, *trp1-1*, *ade2-1*, *leu2-3,112*, *his3-11,15*, *rna15-2*) were based on *rna15-2W* (MATa, *ura3-1*, *trp1-1*, *ade2-1*, *leu2-3,112*, *his3-11,15*, *rna15-2*), kindly provided by Dr F. Lacroute (Centre National de la Recherche Scientifique, Gif-sur-Yvette, France). Correct integration at the *URA3* locus was examined by Southern analysis (17). The yeast strains RH1712 (MATa, *ura3-52::pME800*), carrying the wild-type *GCN4* 3'-UTR in positive orientation, and RH1713 (MATa, *ura3-52::pME801*), with the *GCN4* 3'-UTR inserted in negative orientation in the *in vivo* test system (7), were used as controls. Yeast strains were cultivated on YEPD medium or on SD minimal medium with appropriate supplements at 30°C (18). Yeast *rna15-2W* derivatives were grown at 25°C as the permissive temperature and shifted to 37°C as the restrictive temperature to induce the temperature-sensitive defect of the *rna15-2* allele.

Plasmid construction and cloning

The different point mutations were introduced into the *GCN4* 3'-UTR by site-directed PCR mutagenesis using the oligonucleotides GCN4K-1B (5'-CATTCACTTAGATACGTATT-TAATCTTTTTTATG-3') to generate *GCN4mut-p*, GCN4K-7 (5'-CTATTTTCGTTATGCACGAGAATG-3') to construct *GCN4mut-1*, GCN4K-2 (5'-CTAAGTGAATGTATCCGTA-TCGTTATACACG-3') for *GCN4mut-2*, GCN4K-3 (5'-CTT-ATCTAAGTGAACGTCTCTATTTTCGTTATAC-3') for *GCN4mut-3*, and GCN4K-9 (5'-CATTCACTCAGATAAGAATT-TAATC-3') for *GCN4mut-4*. The oligonucleotides OLCE3 (5'-CGCGGGTACCCTTGAAAATGAGGTTGCCAG-3'), carrying a *KpnI* site, and OLCE6 (5'-GCGCAGATCTCCCC-CATCGTGAGTG-3'), with a *BglII* site, were used as flanking primers. The fragments were cloned into the *KpnI* and *BglII* sites of the multiple cloning site of the test system of plasmid pME798. The resulting plasmids were pME1570 (*GCN4mut-p*), pME1566 (*GCN4mut-1*), pME1567 (*GCN4mut-2*), pME1568 (*GCN4mut-3*) and pME1569 (*GCN4mut-4*). The plasmid pME800 carries the wild-type *GCN4* 3'-UTR in the *in vivo* test system (7).

RNA analysis

Total RNA from *S.cerevisiae* was isolated as described (19). For northern analysis 20 µg of total RNA were separated on a formaldehyde-agarose gel, transferred to a positively charged nylon membrane (Biodyne B; PALL) by electroblotting and hybridized with ³²P-labelled DNA fragments as described previously (7). The DNA fragments were random labelled using the HexaLabel DNA Labelling Kit from MBI Fermentas. Band intensities were visualized by autoradiography and quantified using a BAS-1500 Phosphorimaging scanner (Fuji).

RACE (rapid amplification of cDNA ends)

RACE experiments were performed based on the description in Frohman *et al.* (20). Aliquots of 100 ng of total RNA of the different *S.cerevisiae* strains were reverse transcribed using the M-MuLV reverse transcriptase (MBI Fermentas) with a poly(dT)₁₈ primer. In a following PCR reaction 3'-ends of the different *GCN4* test system transcripts were amplified with the test system-specific primer OLCE9 (5'-GGTCAATCTTTGT-TAAAGAATAGG-3'). OLCE9 binds to an intron sequence which is spliced out in the *ACT1* wild-type mRNA, but which is still present in the transcripts of the *in vivo* test system, where the 3'-splice site is missing. DNA fragments were isolated and cloned into pBluescript SK⁺ (Stratagene). The poly(A) site was determined by sequencing using the T7 Sequencing Kit (Pharmacia).

RESULTS

The *GCN4* 3'-UTR has previously been operationally divided into five different regions, EL1-EL5 (Fig. 1). The different point mutations in this study were located in regions EL2, EL3 and EL4 between the two TTTTAT motifs which were shown to be necessary for efficient 3'-end processing (7). The different alleles of the *GCN4* 3'-UTR were cloned into an *in vivo* test system consisting of the *ACT1* promoter and the *ADHI* 3'-UTR (Fig. 2A). The *ADHI* 3'-UTR is able to process all transcripts which are initiated at the strong *ACT1* promoter.

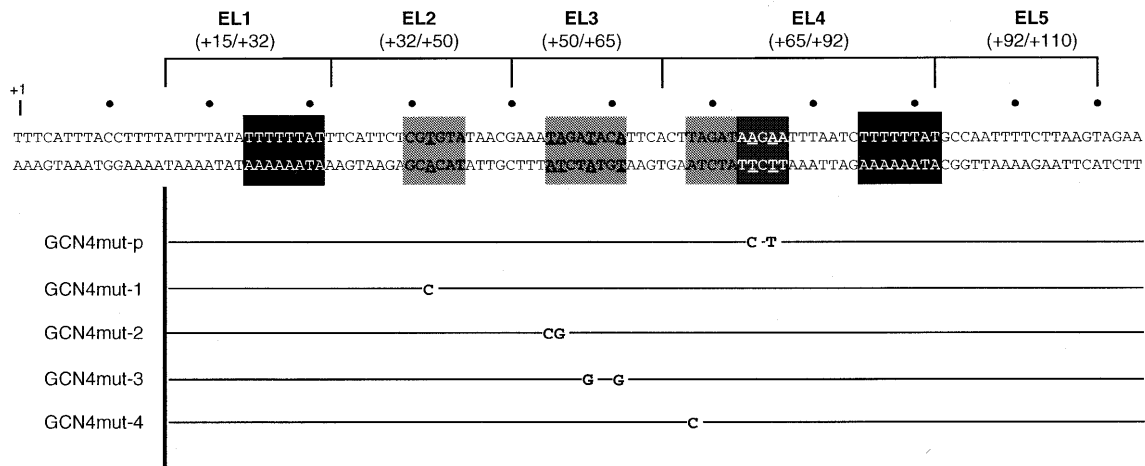


Figure 1. Sequence of the *GCN4* 3'-UTR. The region had previously been divided into five different parts (EL1–EL5) as a result of a deletion analysis (7). The different point mutations which were introduced into the *GCN4* 3'-UTR are shown below. The exchanged nucleotides are underlined in the sequence and the resulting nucleotides of the sense strand are noted. The alleles carrying different point mutations were named GCN4mut-p for alteration of the putative positioning motif AAGAA and GCN4mut-1–GCN4mut-4 for mutations in other sequence motifs according to their 5'→3' succession in the 3'-UTR. The putative sequence motifs in which the point mutations were located are underlined in dark grey for AAGAA and light grey rectangles, respectively. The efficiency motifs of the *GCN4* 3'-UTR were described earlier (7) and are marked by black rectangles. Dark points above the sequence indicate intervals of 10 nt.

In between, the different alleles of the *GCN4* 3'-UTR were cloned (Fig. 2A). By using the *in vivo* test system the 3'-end processing efficiency of the investigated 3'-UTR can be determined. For this purpose the amount of transcripts processed in the *GCN4* 3'-UTR is compared to the total amount of transcripts initiated at the *ACT1* promoter. The test systems containing the different *GCN4* 3'-UTR alleles were integrated into the yeast genome.

Point mutations in an AAGAA motif result in a loss of specificity of poly(A) site selection

The 3'-ends of yeast mRNAs include positioning as well as efficiency elements of mRNA 3'-end processing. Our aim was to analyse sequence motifs in the *GCN4* 3'-UTR for their possible involvement in poly(A) site positioning. AAGAA and similar motifs have been shown to influence selection of the poly(A) site of other yeast genes, e.g. *CYCI* (10). The *GCN4* 3'-UTR contains an AAGAA motif which is located at position +73 relative to the translational stop codon. To investigate its possible role in poly(A) site selection, specific point mutations were introduced resulting in an ACGTA sequence instead of AAGAA (Fig. 1). The allele was named GCN4mut-p (p stands for positioning). The processing efficiency of the GCN4mut-p allele was determined by northern analysis experiments using the 5'-part of the *ACT1* gene as probe. This probe hybridizes to all transcripts of the *in vivo* test system, as well as to the *ACT1* mRNA. The point mutations in AAGAA revealed similar amounts of truncated transcripts processed in the *GCN4* 3'-UTR as the wild-type allele. Therefore, these point mutations had no significant effect on the processing efficiency (Fig. 2B). The wild-type *GCN4* 3'-UTR in negative orientation as an additional control was not functional and revealed a processing efficiency of <5%.

The poly(A) sites of GCN4mut-p were localised by RACE experiments and compared to transcripts containing the wild-type *GCN4* in the *in vivo* test system. The data represent the 3'-ends

of at least 20 different cDNA clones of each construct. The wild-type *GCN4* in positive orientation revealed a single poly(A) site at position +96 relative to the translational stop codon. In contrast, the poly(A) sites of GCN4mut-p were found to form a scattered pattern (Fig. 3). Apart from the wild-type poly(A) site which was predominantly used (43%), additional poly(A) sites were found further downstream at positions +105, +111, +126 and +138. In the case of the poly(A) site at position +138 the preference for PyA as poly(A) site was neglected. The poly(A) tail was added to a C which was formerly followed by another C. The AAGAA motif is therefore a *GCN4* 3'-UTR positioning element which might be responsible for the use of a single specific poly(A) site.

A temperature-sensitive form of the processing factor Rna15p causes an alteration in poly(A) site selection in the wild-type *GCN4* 3'-UTR similar to the AAGAA mutation GCN4mut-p

The 3'-ends of yeast precursor mRNAs are recognised by the processing machinery which is a multisubunit protein complex. The protein factor Rna15p represents a subunit of the processing machinery which directly binds to RNA (14) and which is suggested to be involved in poly(A) site selection (16). Therefore, we wondered how defects in the corresponding *RNA15* gene would affect poly(A) site selection of the *GCN4*-containing test system transcript.

The *in vivo* test system carrying the wild-type *GCN4* 3'-UTR was transformed into a yeast strain carrying an *rna15-2* allele resulting in a temperature-sensitive gene product (14). Total RNA was isolated from cultures grown at the permissive temperature (25°C) and 15, 30 and 60 min after shifting to the non-permissive temperature (37°C). The northern analysis which demonstrates the processing efficiency can be seen in Figure 4. The truncated test system transcript decreased by a factor of three after 60 min incubation at 37°C. Additionally, more read-through transcripts were synthesized and therefore a

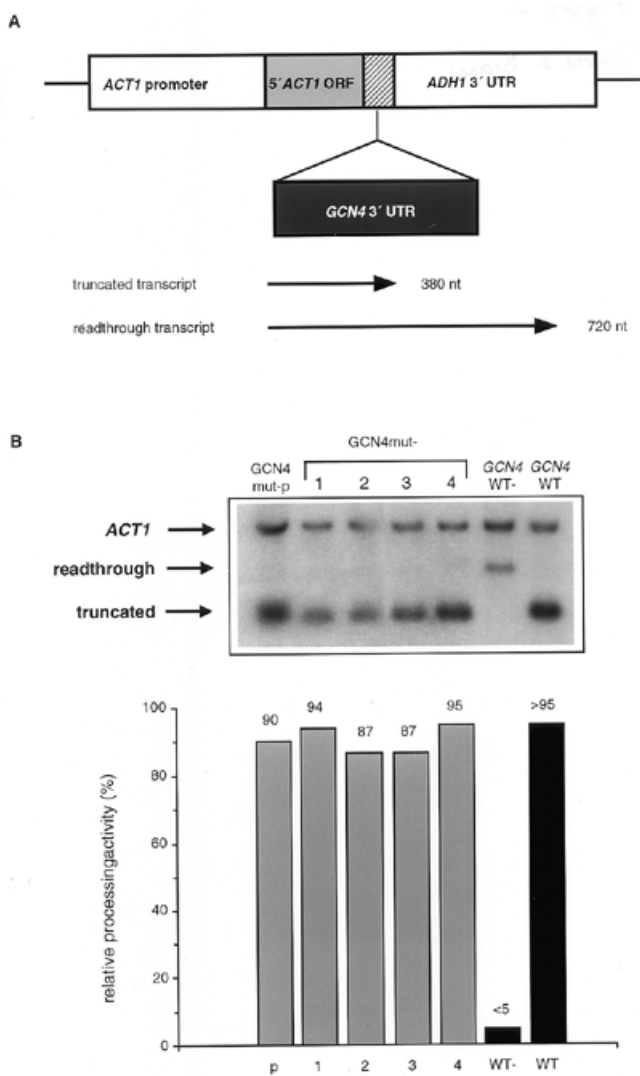


Figure 2. (A) Schematic presentation of the *in vivo* test system which was used to investigate the 3'-end processing activity of the *GCN4* 3'-UTR. It consists of the promoter and a small part of the open reading frame of the *ACT1* gene and the *ADHI* 3'-UTR with a multiple cloning site in between (diagonal stripes) (7). The different *GCN4* 3'-UTRs were cloned into this multiple cloning site. The resulting transcripts are indicated below. The sizes of the read-through and the truncated transcripts are ~720 and 380 nt, respectively. (B) Processing efficiency of the *GCN4* alleles carrying point mutations in the 3'-UTR. The processing efficiency was determined by northern analysis of 20 μ g total RNA, using a 460 bp *MluI*-*KpnI* fragment of the 5'-region of the *ACT1* gene as probe. The upper band of the resulting three bands represents the transcript of the genomic *ACT1* gene. The band in the middle shows a read-through transcript which is processed in the *ADHI* 3'-UTR of the test construct. The lower band represents truncated transcripts of the test constructs which were processed in the *GCN4* 3'-UTR. The wild-type *GCN4* 3'-UTR in positive, functional or negative, non-functional orientation served as controls. Quantification was performed by phosphorimager analysis. The processing activities are expressed by the percentage of truncated transcripts compared to the sum of truncated and read-through transcripts. The columns represent northern hybridizations of at least three independent RNA isolations. The standard deviations did not exceed 15%.

change in the ratio between read-through and truncated transcripts can be observed. After 60 min incubation at the non-permissive temperature the processing activity of the *GCN4* 3'-UTR is decreased to 81 compared to 97% at the

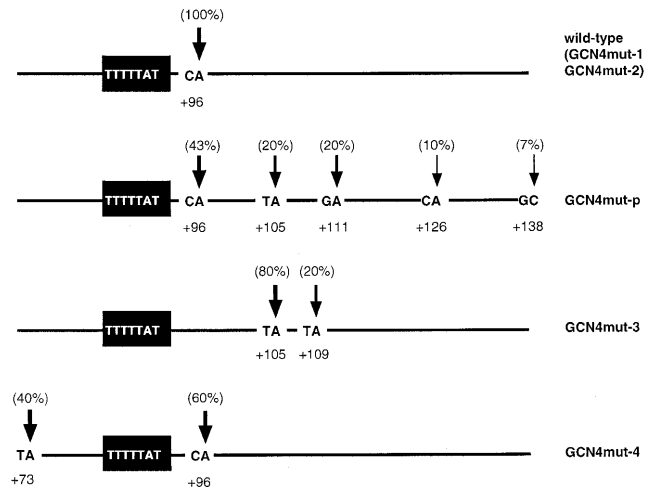


Figure 3. Schematic presentation of the poly(A) site selection of the wild-type *GCN4* 3'-UTR and the alleles carrying different point mutations. The poly(A) sites were determined by RACE experiments. The TTTTAT sequence corresponds to the efficiency motif located more downstream. The poly(A) sites are indicated by arrows, as are the dinucleotides where cleavage occurred. The positions are relative to the translational stop codon. The frequency with which the poly(A) site was chosen is shown in parentheses and indicated by the thickness of the arrow. Each arrow represents at least two independent cDNA clones. Altogether 20 cDNA clones of each test construct were sequenced, with the exception of the *GCN4mut-p* allele, for which 30 clones were used.

permissive temperature. Rna15p seems to have an auxiliary function in efficient processing of the precursor mRNA.

To determine alterations in poly(A) site selection, RACE experiments were performed for the non-shifted culture and the culture which was grown for 60 min at 37°C, respectively. The test system RNA of the non-shifted culture exhibits a single poly(A) site at position +96 relative to the translational stop codon, which is identical to the wild-type yeast strain carrying an intact *RNA15* allele (Fig. 3). However, the transcripts of the temperature-shifted cells revealed additional poly(A) sites further downstream (Fig. 5A). The major poly(A) site was located at the wild-type position (+96) and was chosen in 60% of all investigated test system cDNAs. Additional poly(A) sites were located at positions +102, +105, +111, +132 and +137, respectively. Each of these poly(A) sites was used at a much lower frequency than the poly(A) site at position +96. The preference for PyA is not strictly followed since CT and GC were also chosen as poly(A) sites. The positions of the new poly(A) sites of the intact *GCN4* 3'-UTR in an *rna15* mutant strain were similar to the pattern and the reduced sequence specificity which were obtained by the point mutated AAGAA motif in a wild-type *RNA15* genetic background. These data suggest that Rna15p interacts with the AAGAA motif at position +73 of the *GCN4* 3'-UTR.

The *GCN4* 3'-end comprises additional positioning motifs for poly(A) sites which resemble putative efficiency elements

Changes in the AAGAA motif or function of the Rna15 protein defocus 40–60% of the 3'-ends of the *GCN4* mRNA. The remaining 3'-ends are still polyadenylated at the same position as in the presence of the wild-type *GCN4* sequence and intact Rna15p. Therefore, we expected that additional positioning

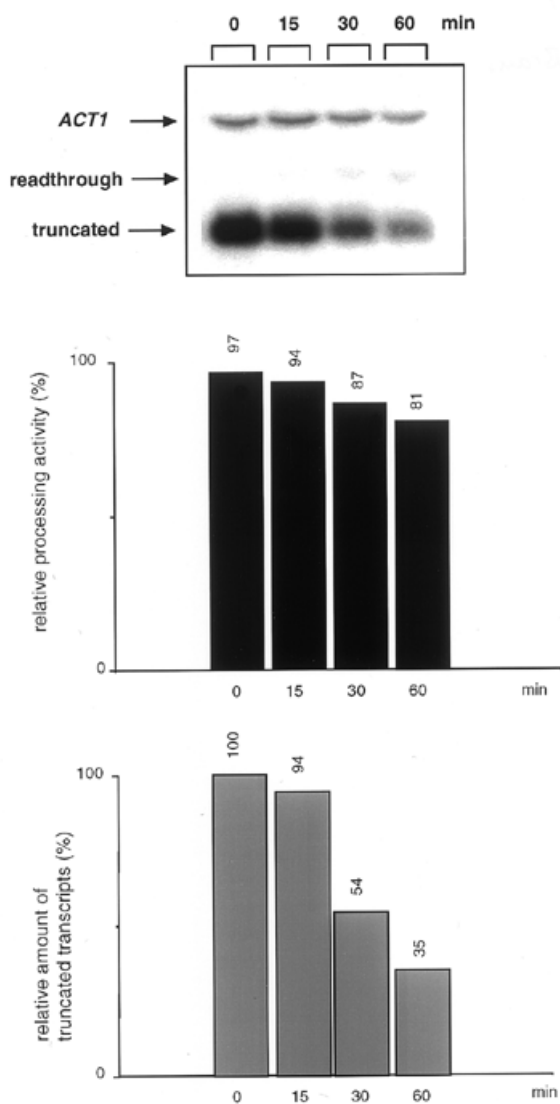


Figure 4. Processing activity of the test construct carrying a wild-type *GCN4* 3'-UTR in a strain with the *ma15-2* temperature-sensitive mutant allele. The cultures were grown at 25°C and then shifted to 37°C. Total RNA was isolated prior to the temperature shift (0) and 15, 30 and 60 min after it. Aliquots of 20 µg of RNA from each sample equalized by the extinction at a wavelength of 260 nm were hybridized with a probe from the 5'-part of the *ACT1* gene as described in Figure 2B. The processing activities show the amount of truncated transcripts compared to the total amount of all transcripts expressed by the *in vivo* test system. Quantification of the decrease in total amounts of truncated transcripts is shown below. The amount of truncated transcripts prior to the temperature shift (0 min) was set as 100%. The columns represent an average of three independent measurements and the standard deviation did not exceed 15%.

motifs might be present in the *GCN4* 3'-UTR. A single point mutation in a TAGT motif changed efficiency of 3'-end processing as well as poly(A) site selection of the *TRP4* 3'-UTR (12). In between the two TTTTAT motifs of the *GCN4* 3'-UTR, which are necessary for efficient 3'-end processing (7), there are different variations of TA(T)GT and TATATA motifs, which are the common efficiency motifs in many yeast genes. Different point mutations were introduced into these sequence

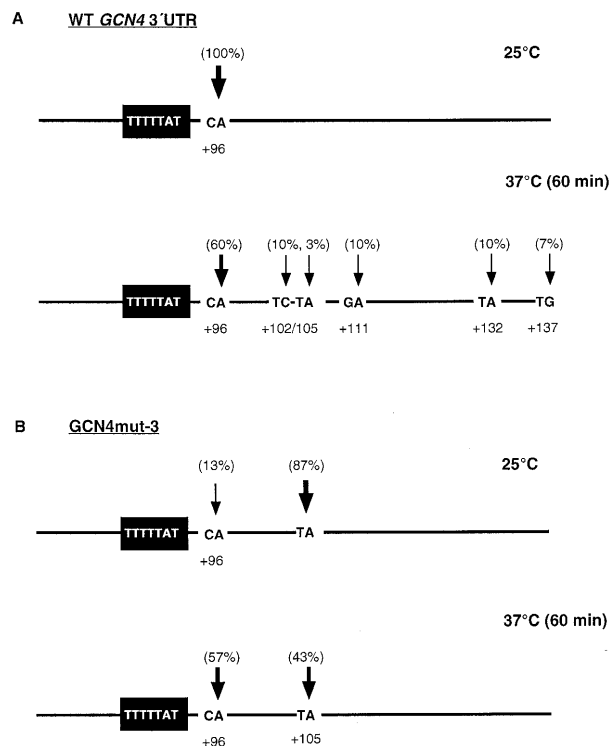


Figure 5. (A) Schematic presentation of poly(A) site selection of transcripts resulting from the *in vivo* test system carrying the wild-type *GCN4* 3'-UTR in an *ma15-2* strain. The poly(A) sites were detected by RACE experiments of RNA from cultures grown at the permissive temperature (25°C) and cultures grown for 60 min at the restrictive temperature (37°C). Poly(A) sites are indicated by arrows, as are the dinucleotide sequences where cleavage occurs. The position of the poly(A) site relative to the translational stop codon is shown below and the relative frequency of these mRNA ends in parentheses above the arrows. The data represent 30 independent clones of three independent RNA isolations for each time point. (B) Schematic presentation of poly(A) sites which were used in the *GCN4mut-3* allele in an *ma15-2* background. Determination and presentation of poly(A) sites correspond to (A).

motifs to determine if there is any influence of these degenerate motifs on poly(A) site selection. Selected sequence motifs in the *GCN4* 3'-UTR were TGTATA, starting at position +42 relative to the translational stop codon, TAGATACA at position +54 and TAGAT at position +68. The point mutations were numbered according to their 5'→3' appearance in the *GCN4* 3'-UTR resulting in *GCN4mut-1* for the point mutation in TGTATA. *GCN4mut-2* and *GCN4mut-3* were located in the TAGATACA motif and *GCN4mut-4* in the TAGAT motif.

Our first objective was to test by northern analysis whether one of these point mutations, which were inserted in putative efficiency motifs, changes the rate of RNA processing (Fig. 2B). However, the efficiency of 3'-end processing was not significantly affected by any of the point mutations in the sequence motifs. They all showed >85% transcripts which were processed in the *GCN4* 3'-UTR. Therefore, these sequences are not significantly involved in efficient RNA precursor processing.

To determine the poly(A) sites of the different *GCN4* test system transcripts RACE experiments were performed (Fig. 3). *GCN4mut-1* and *GCN4mut-2* revealed only the wild-type

poly(A) site at position +96. These two alterations of the *GCN4* 3'-UTR therefore had an effect neither on the selection of the poly(A) site nor on the efficiency of 3'-end processing.

The point mutation *GCN4mut-3* changing the TAGATACA motif to TAGAGACG revealed two poly(A) sites located further downstream (Fig. 3). The wild-type poly(A) site was not used in this mutated allele. The major poly(A) site of this construct which occurred in 80% of all sequenced RACE products is located at a TA motif at position +105. A less frequently chosen poly(A) site was found at another TA motif at position +109. Therefore, this element seems to be important for poly(A) site selection at the wild-type site. Alternatively, a new positioning motif could have been created by this point mutation which directs cleavage at the downstream poly(A) site in a very efficient way.

In contrast, the point mutation *GCN4mut-4* concerning the TAGAT motif adjacent to the putative positioning motif AAGAA resulted in the wild-type poly(A) site and an additional poly(A) site which was located further upstream at a TA motif at position +73 (Fig. 3). They were chosen with comparable frequency, the wild-type poly(A) site in 60% and the upstream poly(A) site in 40% of all cases. Processing at the novel poly(A) site separates the mutated TAGAT sequence from the AAGAA positioning motif which is located immediately downstream and adjacent to it. Therefore, changes in this element result in 3'-ends which are located at the positioning element.

Expression of the *GCN4mut-3* allele in an *rna15-2* strain results in a novel pattern of selected poly(A) sites

The *GCN4mut-3* allele was transformed into the *rna15-2*-carrying yeast strain to investigate the effect of non-functional Rna15p on poly(A) site selection. Test system transcripts isolated at the permissive temperature and up to 60 min after the shift to the restrictive temperature revealed similar amounts of transcripts as the wild-type *GCN4* allele (data not shown). RACE experiments were performed on RNA isolated prior to and 60 min after incubation at 37°C (Fig. 5B). The downstream poly(A) site at position +105, which was also predominantly used in a wild-type *RNA15* background, was found in 87% of all sequenced cDNAs in the culture grown at the permissive temperature. In contrast to the wild-type strain, the *rna15* mutant strain occasionally showed the wild-type poly(A) site at position +96 even at the permissive temperature (13% of all cases). Temperature shift to 37°C resulted in an altered percentage of the use of wild-type compared to the downstream poly(A) site. After 60 min at 37°C the wild-type poly(A) site was preferred in 57% of all sequenced cDNAs, whereas use of the downstream poly(A) site was reduced to 43%. There were no further downstream poly(A) sites detectable, as was the case for the wild-type *GCN4* allele in the *rna15* background. Whereas the *GCN4mut-3* mutation alone is unable to use the wild-type poly(A) site, the additional defect in the *RNA15* gene suppresses this effect and allows the use of this site again. The novel pattern of poly(A) site selection therefore reveals some of the features of each single mutant strain. Possible explanations are additional, yet unknown RNA-protein interactions between the pre-mRNA and the processing machinery.

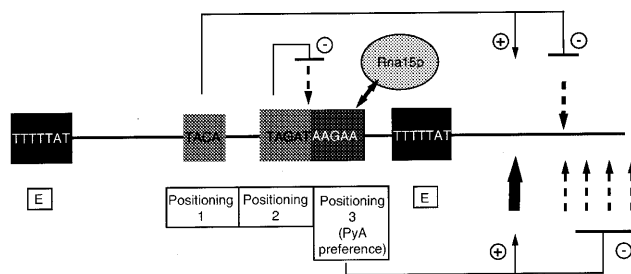


Figure 6. Model of the possible action of three different positioning motifs in the *GCN4* 3'-UTR. The efficiency motifs (E) are underlined with black rectangles and the positioning motifs with light grey and dark grey rectangles, respectively. The wild-type poly(A) site is marked by a thick arrow. Use of the wild-type poly(A) site is enhanced by positioning motifs 1 and 3 (+). Cryptic poly(A) sites which were detected in the case of point mutated positioning motifs are indicated by dashed arrows. Since they are only used when specific positioning motifs are altered these motifs might normally prevent their use (-). Rna15p is suggested to interact with positioning motif 3 due to similarities in the poly(A) site selection in both a point mutated AAGAA motif and a wild-type *GCN4* 3'-UTR in an *rna15* mutant strain.

DISCUSSION

In this report we have shown that three different positioning motifs in the *GCN4* 3'-UTR influence poly(A) site selection in three different ways. The arrangement of these positioning motifs and their effects when point mutated are shown as a model in Figure 6. The most upstream located positioning motif is a TACA sequence which when point mutated results in completely changed positioning of the poly(A) sites. Poly(A) sites were exclusively shifted downstream of the wild-type poly(A) site. The change to the new poly(A) sites is very specific, since only one major and a minor poly(A) site nearby could be detected. In the wild-type 3'-UTR these poly(A) sites are completely cryptic and were never detected. Therefore, this positioning motif seems to prevent usage of these cryptic sites. It cannot be ruled out that by introducing the point mutation a new positioning motif was created which directs the cleavage machinery to the downstream poly(A) site. However, the changes result in a more GC-rich motif, making it unlikely that it functions as a processing element, which are mostly AT-rich.

The second positioning motif, TAGAT, is located close to the third, the AAGAA positioning motif. In the case of the TAGAT motif the wild-type poly(A) site is still used in approximately half of all cases but an additional upstream poly(A) site turns up. Remarkably, this new poly(A) site is located directly in between the TAGAT and the AAGAA motifs. As a consequence the positioning motif AAGAA is located downstream of the poly(A) site on the 3' cleavage product of the precursor mRNA. This is reminiscent of poly(A) site determinants of higher eukaryotes which include sequence motifs located upstream and downstream of the poly(A) site. The upstream motif is the highly conserved AAUAAA motif and the downstream motif is U- or GU-rich. A certain distance between these two motifs is necessary in higher eukaryotes to specify the poly(A) site (21). In yeast, the altered TAGAT motif might partially inactivate the adjacent AAGAA motif

because the newly created poly(A) site concerns the first A of the AAGAA motif. These specific point mutations in the TAGAT sequence might give access to the upstream poly(A) site which was formerly hidden, e.g. by processing factors. It implies that this region must then be accessible to cleavage factors.

The role of the AAGAA motif differs from those of the other positioning motifs in the *GCN4* 3'-UTR. Point mutations in AAGAA result in a scattered pattern of poly(A) sites. The wild-type poly(A) site was still predominantly used but several additional poly(A) sites occurred downstream of it. None of these cryptic poly(A) sites was detected in the wild-type *GCN4*, suggesting that an intact AAGAA motif represses their use. Additionally, AAGAA might be necessary to keep the preference for a PyA dinucleotide as selected poly(A) site. The PyA preference is well established in many yeast genes. Point mutations in PyA motifs of poly(A) sites of different genes lead to the predominant selection of other PyA sites nearby (22).

The distance between the positioning motif and poly(A) site might be an important determinant of poly(A) site selection. A computer alignment of yeast 3'-UTRs revealed a significant peak of A-rich positioning motifs at a distance of ~20–25 nt from the poly(A) site (3). The AAGAA sequence of the *GCN4* 3'-UTR exactly fits this model.

In higher eukaryotes, selection of the specific poly(A) site requires a minimal distance between the upstream and the downstream processing sequence motifs. It was suggested that binding of processing factors is the reason for the spacing requirement (21). For both the upstream and the downstream motifs, binding of specific processing factors has been demonstrated. The 160 kDa subunit of the cleavage polyadenylation specificity factor CPSF-160 binds to AAUAAA (23) and the 64 kDa subunit of the cleavage stimulation factor CstF-64 to the downstream motif. CstF-64 *in vitro* crosslinking experiments showed that the location of CstF-64 is important for poly(A) site selection (24).

In this report we suggest that the Rna15p processing factor, which has high similarity to CstF-64, especially in its RNA-binding domain (25), interacts with the AAGAA motif. This assumption is based on the finding that a point mutated AAGAA motif results in the same pattern of poly(A) sites as a wild-type *GCN4* 3'-UTR in a strain carrying a temperature-sensitive allele of *RNA15*. In both cases, the poly(A) sites were dispersed over a broader region. A connection between the positioning motif and the Rna15p processing factor is likely; however, a direct interaction remains to be proven. SELEX experiments of Rna15p revealed a UGC/UGUAUUC/UUCC sequence as the RNA consensus motif (15). SELEX studies were also done with the homolog of Rna15p in higher eukaryotes, CstF-64. Although the RNA-binding domains are very similar, SELEX studies performed independently by two laboratories revealed different consensus sequences. In one case this was an AUGCGUUCUCGUCC motif (26), in the other GU-rich sequences without strict consensus (15). Neither of these consensus sequences resembles the AAGAA motif. There is more similarity to the other positioning sequences of the *GCN4* 3'-UTR. However, SELEX studies provide *in vitro* data which neglect the fact that processing factors tend to bind cooperatively (23).

Point mutations in the first two positioning motifs allow the specific use of other poly(A) sites in the presence of an intact AAGAA motif as putative interaction site for Rna15p. These point mutations do not change the predominance of one or a maximum of two poly(A) sites, in contrast to the multiple sites which are a result of defects in the *RNA15* gene. Even at the restrictive temperature, the *rna15* mutant strain showed a preference for two poly(A) sites of the *GCN4*mut-3 allele. In this case defective Rna15p led only to the additional use of the wild-type poly(A) site which could not be detected in a strain carrying wild-type *RNA15* and the *GCN4*mut-3 allele. The combination of mutations in *RNA15* and a sequence motif in the *GCN4* 3'-UTR therefore results in a novel phenotype which shares some of the features of each single mutant strain. Deletion of a 15 bp spanning region of the *GCN4* 3'-UTR performed in a former study (7) revealed that the wild-type poly(A) site and two additional downstream poly(A) sites served with equal frequency. Destruction of a whole region could therefore concern different positioning motifs and the effect need not be additive but could result in a different pattern of poly(A) site selection.

The point mutations in the *GCN4* 3'-UTR might also have altered the RNA secondary structure, providing different access to the cleavage sites in the 3'-UTR. According to an MFOLD analysis (based on 27) the *GCN4* 3'-UTR does not fold into any stable secondary structure, which makes it unlikely that the effect of point mutations in this region would only be due to alterations of the secondary structure.

Alternatively, these sites could be targets for other proteins of the processing machinery. There are several protein factors in yeast which have been shown to possess RNA-binding activity. A possible candidate could be the Hrp1/Nab4 protein. It is part of the CFI complex, like Rna15p, and it was characterized by two different laboratories (28,29). Hrp1p/Nab4p was shown to bind (UA)_n motifs (30,31). TATATA motifs mostly function as efficiency motifs of mRNA 3'-end formation. The involvement of TATATA-binding factors in poly(A) site selection remains speculative, at least in those 3'-UTRs in which efficiency and positioning motifs are independent of each other, like the *GCN4* 3'-UTR. However, depletion of Hrp1p/Nab4p resulted in the use of specific upstream poly(A) sites of *CYC1* and the *GAL7* mRNA (29). The wild-type poly(A) site was still used but with lower frequency. The effects of a *hrp1/nab4* mutation on poly(A) site selection therefore resemble the effects obtained by point mutations in the TAGAT motif upstream of the AAGAA motif. The involvement of Hrp1p/Nab4p in poly(A) site selection of the *GCN4* 3'-UTR remains to be proven.

In summary, our data suggest more contact points between the 3'-end processing machinery and the precursor mRNA in yeast than previously assumed.

ACKNOWLEDGEMENTS

We thank Sabine Eckert and Ralph Pries for critical reading and helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft, the Volkswagen-Stiftung and the Fonds der Chemischen Industrie.

REFERENCES

1. Zhao, J., Hyman, L. and Moore, C. (1999) *Microbiol. Mol. Biol. Rev.*, **63**, 405–445.
2. Wahle, E. and Rügsegger, U. (1999) *FEMS Microbiol. Rev.*, **23**, 277–295.
3. Graber, J.H., Cantor, C.R., Mohr, S.C. and Smith, T.F. (1999) *Nucleic Acids Res.*, **27**, 888–894.
4. Guo, Z., Russo, P., Yun, D.-F., Butler, J.S. and Sherman, F. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 5983–5990.
5. Guo, Z. and Sherman, F. (1996) *Mol. Cell. Biol.*, **16**, 2772–2776.
6. Hyman, L.E., Seiler, S.H., Whoriskey, J. and Moore, C.L. (1991) *Mol. Cell. Biol.*, **11**, 2004–2012.
7. Egli, C.M., Springer, C. and Braus, G.H. (1995) *Mol. Cell. Biol.*, **15**, 2466–2473.
8. Egli, C.M., Düvel, K., Trabesinger-Rüf, N., Irniger, S. and Braus, G.H. (1997) *Nucleic Acids Res.*, **25**, 417–422.
9. Aranda, A., Pérez-Ortín, J.E., Moore, C. and del Olmo, M. (1998) *RNA*, **4**, 303–318.
10. Russo, P., Li, W.Z., Guo, Z. and Sherman, F. (1993) *Mol. Cell. Biol.*, **12**, 7836–7849.
11. Guo, Z. and Sherman, F. (1995) *Mol. Cell. Biol.*, **15**, 5983–5990.
12. Düvel, K., Egli, C.M. and Braus, G.H. (1999) *Nucleic Acids Res.*, **27**, 1289–1295.
13. Keller, W. and Minvielle-Sebastia, L. (1997) *Curr. Opin. Cell. Biol.*, **9**, 329–336.
14. Minvielle-Sebastia, L., Winsor, B., Bonneaud, N. and Lacroute, F.M. (1991) *Mol. Cell. Biol.*, **11**, 3075–3087.
15. Takagaki, Y. and Manley, J.L. (1997) *Mol. Cell. Biol.*, **17**, 3907–3914.
16. Mandart, E. and Parker, R. (1995) *Mol. Cell. Biol.*, **15**, 6979–6986.
17. Southern, E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
18. Rose, M.D., Winston, F. and Hieter, P. (1990) *Methods in Yeast Genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Cross, F.R. and Tinkelenberg, A.H. (1991) *Cell*, **65**, 875–883.
20. Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.
21. Chen, F., MacDonald, C. and Wilusz, J. (1995) *Nucleic Acids Res.*, **23**, 2614–2620.
22. Heidmann, S., Obermaier, B., Vogel, K. and Domdey, H. (1992) *Mol. Cell. Biol.*, **12**, 4215–4229.
23. Murthy, K.G.K. and Manley, J.L. (1995) *Genes Dev.*, **9**, 2672–2683.
24. MacDonald, C.C., Wilusz, J. and Shenk, T. (1994) *Mol. Cell. Biol.*, **14**, 6647–6654.
25. Kessler, M.M., Zhao, J. and Moore, C.L. (1996) *J. Biol. Chem.*, **271**, 27167–27175.
26. Beyer, K., Dandekar, T. and Keller, W. (1997) *J. Biol. Chem.*, **272**, 26769–26779.
27. Zuker, M. and Stiegler, P. (1981) *Nucleic Acids Res.*, **9**, 133–148.
28. Kessler, M.M., Henry, M.F., Shen, E., Zhao, J., Gross, S., Silver, P.A. and Moore, C.L. (1997) *Genes Dev.*, **11**, 2545–2556.
29. Minvielle-Sebastia, L., Beyer, K., Krecic, A.M., Hector, R.E., Swanson, M.S. and Keller, W. (1998) *EMBO J.*, **24**, 7454–7468.
30. Zhao, J., Kessler, M.M. and Moore, C.L. (1997) *J. Biol. Chem.*, **272**, 10831–10838.
31. Chen, S. and Hyman, L.E. (1998) *Nucleic Acids Res.*, **26**, 4965–4974.