Chapter 3

Transcription Termination of \textit{Am41} Genomic Clone in \textit{Schizosaccharomyces pombe} and Delineation of Different \textit{cis} Elements Involved in the Process

3.1. Introduction

Transcriptional termination constitutes the basic premise of gene expression on which relies the efficient operation of a variety of cellular processes ranging from correct gene expression to homologous recombination and chromosome transmission. Transcription termination, defined as the dissociation of the ternary complex into its constituent parts, occurs further downstream of the matured mRNA end and is required to release the polymerase for subsequent rounds of transcription as well as to prevent transcriptional interference. Transcriptional interference arises when an elongating polymerase fails to terminate transcription and reads on into a co-transcribed downstream transcription unit. This results in reduced expression of the downstream gene through disruption of interactions at the promoter (Hansen \textit{et al.}, 1998). Besides promoter occlusion, potential problems could also arise when two promoters face each other in the event of convergent transcription, leading to the formation of RNA duplexes. In yeast, transcriptional interference is believed to be of particular significance due to the compact nature of the yeast genome (Oliver \textit{et al.}, 1992). Furthermore, yeast autonomously replicating sequences (ARSs) and centromere (CEN) elements, both of which are required for chromosome maintenance, as well as sites of meiotic gene conversion, are also sensitive to transcriptional interference. In addition, premature termination is important as a regulatory mechanism for the expression of several cellular and viral genes (Hyman and Moore, 1993).
There is an inherent dichotomy of elongation versus termination: Once transcription is under way, RNA polymerase grasps the template very securely to synthesize long RNA molecules rapidly and without error or premature arrest; yet it must stop and release with high efficiency at the designated end. Termination is an irreversible process, encompassing both release of the transcript and dissociation of polymerase from the template. Termination sites themselves comprise of both pause and release elements, and separable recognition elements when additional factors also participate. A pause refers to transient hesitation of the elongation complex, which is capable of fully resuming transcription along the template after a dwell time of some characteristic half-life. On the other hand, a release refers to release of the nascent transcript and dissociation of the enzyme from the template. The juxtaposition of separate components echoes a familiar theme in biological systems, ensuring that the accidental occurrence of single element does not trigger an unwanted event. Although of remarkable variety, termination sites fall naturally into two classes, intrinsic or factor-dependent, as alternative default states of 'on' or 'off', respectively. RNA structure plays an important role in the function of many intrinsic sites, and antagonistic effects often interfere with formation or recognition of such structure. In other cases, mutually exclusive interactions between different regions of a highly structured transcript can provide conditional transcription termination in response to cellular factors.

In the milieu of the cell, the transcription rate was long thought to be constant, carried out by a fixed elongation-specific conformation of the enzyme that was acquired during synthesis of the first few dozen nucleotides. This stable elongation complex then continued transcription with phenomenal processivity. Indeed, polymerase molecules are remarkably resistant to many factors that might disrupt their progress and hence, unusual cellular strategies for transcription termination and its control must be employed. A clear understanding of such mechanisms is essential to illuminate polymerase functions. At the heart of these mechanisms is the structure of the elongating RNA polymerase molecule and its associated factors, in relation to the DNA template and the growing RNA chain. Each of several regions is a potential locus for control (Platt, 1998).
Obviously, the variety of polymerase increases the complexity of the entire process of transcription termination. However, closer investigation and a rather clear understanding of the signals that direct formation of complexes involved in the process is facilitated in situations when the transcripts are synthesized at a higher order of magnitude mediated by RNA pol-I and pol-III. The high level of the transcripts synthesized by RNA pol-I and pol-III has facilitated the investigation of signals that direct their formation. RNA pol-I synthesizes ribosomal RNA as a long precursor whose 3'-end results directly from a termination event. In *S. cerevisiae*, pol-I transcription terminates some 10-20 bp upstream of the binding site for a sequence specific protein Reb1p, which mediates polymerase pausing (Jeong *et al.*, 1995). An additional T-rich element located upstream of the protein binding site is also required to pause and release the polymerase. Thus in case of pol-I, termination depends on a pause signal, a release element in the template, and a DNA-binding protein – the basic mechanism appears conserved throughout eukaryotes (Lang *et al.*, 1994). The pause signal is the binding site for the protein to a species-specific consensus motif in the DNA that lies just downstream from the release element. On the other hand, pol-III is capable of mediating transcriptional termination unassisted, with four regions of the second largest subunit of the yeast pol-III being implicated in affecting termination efficiency. A string of T residues preferentially embedded in a GC-rich region on the non-template strand constitute an RNA pol-III termination signal, which causes significant pausing of the polymerase followed by release from the template (Matsuzaki *et al.*, 1994).

In contrast, far less is known about the process of RNA pol-II transcription termination, partly due to the instability of the primary transcript and lower transcript levels compared with genes transcribed by pol-I and pol-III (Hansen *et al.*, 1998). Prior to a discussion on the knowledge available pertaining to transcription termination by RNA pol-II, which transcribes eukaryotic mRNA, it will be worthwhile to have a glimpse of how 3'-end processing event of this RNA species takes place. It is important precisely because a range of studies revealed that the eukaryotic polyadenylation signals at the 3'-end of pre-mRNA are required for termination of transcription in higher eukaryotes, budding yeast and fission yeast (Birse *et al.*, 1998).
Biogenesis of functional eukaryotic mRNAs requires the addition of a poly(A) tail at their 3'-ends. The only known exceptions are the major histone mRNAs in metazoan organisms. 3'-end formation plays a role in many aspects of mRNA metabolism (Barabino and Keller, 1999). The 3'-end of most eukaryotic mRNA is formed by coupled endonucleolytic cleavage of the nascent transcript synthesized by pol-II followed by polyadenylation (addition of ~200 As) of the upstream cleavage product. In the cytoplasm, the poly (A) tail on the message plays a role in stability and translatability and stabilizes RNA from degradation by preventing association with the degradation machinery. While initiation of RNA polymerase II transcripts is an important starting point in gene expression, the job is not done until the poly (A) tail is added and the mRNA is exported and translated in the cytoplasm. Transcription ends well only when the message ends well; therefore, polyadenylation is an important means to the end of the message (Edwalds-Gilbert et al., 1997 and references therein).

The signals directing cleavage/polyadenylation have been studied extensively in both vertebrates and the yeast S. cerevisiae and are now well characterized (Guo and Sherman, 1996; Wahle and Keller, 1996). In vertebrates, three sequence elements determine the precise site of 3'-end cleavage and polyadenylation of pre-mRNAs. An invariant poly(A) signal, the hexanucleotide AAUAAA, is present in the 3'-untranslated region (3'-UTR), near the mature mRNA end of many genes (~80%), while AUUAAA is found less frequently. Cleavage occurs at the poly(A) site, ~11-23 nucleotide downstream of the hexanucleotide. The third element is a GU- or U-rich region, usually 10-30 bases downstream of the cleavage site. The AAUAAA or some variant of it in the downstream region and their relative positions define the approximate site at which cleavage will occur for most poly(A) sites; secondary structure can shift the site of poly(A) tail addition slightly (Edwalds-Gilbert et al., 1997). In some viral genes there is an additional element upstream of the hexanucleotide which can aid in efficient poly(A) site recognition. The efficiency of 3'-end processing can be increased by sequence elements located upstream of the polyadenylation signal. Again, many eukaryotic and viral genes give rise to mRNAs that differ in their 3'-ends due to the choice between alternative poly(A) sites. Therefore, regulated 3'-end processing of specific RNAs could potentially be used to modulate gene expression in a tissue or developmental stage-specific
manner. Yeast polyadenylation signals are often degenerate and redundant to a higher degree than that found in mammals (Minvielle-Sebastia and Keller, 1999). The more efficient a poly (A) site is at processing in vitro, the more efficient it is at generating termination-competent RNA polymerase II elongation complexes and mature RNA. Poly (A) site strength can directly influence the amount of cytoplasmic RNA produced from a transcript (Edwalds-Gilbert et al., 1997 and references therein); therefore, changing polyadenylation efficiencies can have a profound effect on the amount and nature of a gene product. Efficient transcription termination signals must exist to prevent such interference and the termination event must take place fairly close to the poly(A) site. Indeed, evidence suggests that transcription termination occurs within approximately 100 nt of yeast poly(A) sites (Russo and Sherman, 1989).

Far less is known about mRNA 3′-end processing in plants. The discovery of the importance of the AAUAAA motif for polyadenylation in animal systems naturally prompted a search for similar motifs in plant genes. It soon became apparent that AAUAAA is not universally conserved as a poly (A) signal; analysis of plant 3′-UTRs found an exact match of this sequence at an appropriate position in less than 40% of cases (Hunt et al., 1987; Joshi, 1987). In addition to sequence analyses, experimental data showed that animal poly (A) signals were not properly recognized in plant cells, again suggesting functional differences in the cis-acting sequences controlling 3′-end formation. Another distinguishing feature is that animal genes normally have a single poly (A) site, whereas in plants the position of cleavage can be quite heterogeneous within a single transcription unit, leading to the production of mRNA populations with a variety of end points (Dean et al., 1986; Klahre et al., 1995). All of these features suggest that the processes of mRNA cleavage and polyadenylation in plants might differ mechanistically from those characterized in vertebrate systems.

While substantial progress has been made in the characterization of the basic 3′-end processing apparatus, 3′-end formation has long been considered a fundamental but invariant process and its regulatory aspects are still only poorly understood. Surprisingly enough, this relatively simple event of pre-mRNA 3′-end processing, requires many trans-acting protein factors. In vertebrates, the cleavage stimulation factor (CstF) and the cleavage factors CF Im
and CF IIₘ participate only in the cleavage reaction. Poly(A) polymerase (PAP) and the cleavage and polyadenylation specificity factor (CPSF) participate in both cleavage and polyadenylation. Finally, the poly(A) binding protein II (PAB II) increases the efficiency of polyadenylation and specifies the correct length of the poly(A) tail (Wahle and Ruegsegger, 1999). The choice of the poly(A) site has been shown to be influenced by, (1) the intrinsic strength of the cis-acting sequence elements that define the cleavage site, (2) changes in the concentration or the activity of constitutive polyadenylation factors, and (3) the expression of stage-specific regulatory factors (Reviewed by Barabino and Keller, 1999).

The interplay of these diverse and distinct elements and factors behind the RNA pol-II mediated 3'-end processing event contributes towards the process in a manner which is far too complex and critical. Termination of transcription by RNA pol-II, unlike the well characterized termination events in eukaryotic genes transcribed by RNA pol-I and pol-III, remains obscure. A bipartite signal has been proposed to direct pol-II termination in higher eukaryotes though the termination sites display great variety and complexity. Termination of full-length precursor per se is often considerably beyond the mature 3'-end of the transcript, which itself is generated by post-transcriptional cleavage and addition of a poly(A) tail. This terminal maturation event and transcription termination by pol-II appear to be interdependent, even though the latter may transcribe many kilobases beyond the polyadenylation site before termination occurs. A functional polyadenylation element acts as the upstream signal and a prerequisite of transcriptional termination with mutation of this highly conserved processing signal leading to reduced termination efficiency downstream (Connelly and Manley, 1988, Logan et al., 1987). Further, a strong poly(A) site transposed upstream of a weak one can improve distal transcription termination. Moreover the strength of this processing signal, correlates with termination efficiency (Edwalds-Gilbert et al., 1993). The nature of the downstream signal, located in the region where pol-II termination takes place, remains poorly defined. Various types of downstream element (DSE) including a distorted DNA template, trans-acting factors binding to the DNA and the structure of the transcript have been implicated in mediating pol-II termination (Kerppola and Kane, 1990; Connelly and Manley, 1989; Ashfield et al., 1994; Pribyl and Martinson, 1988). One of the functions of DSE may be to pause the elongating polymerase by perturbing the normal mode
of elongation in a similar fashion to that observed for both pol-I and pol-III (Enriquez-Harris et al., 1991; Eggermont and Proudfoot, 1993).

Analysis of pol-II transcriptional termination is complicated by the upstream mRNA 3'-end processing events, cleavage and polyadenylation, and the observation that termination sites appear to be heterologous (Reviewed by Colgan and Manley, 1997). A series of studies has enabled the beginning of an understanding of the molecular events which lead to transcription termination by RNA pol-II. It is known that the enzyme is capable of transcribing a gene well beyond the poly(A) signal. Termination regions mapped for the genes transcribed by pol-II lie between 100 and 4000 bp downstream of the poly(A) site and is imprecise in nature (Reviewed by Proudfoot, 1989). Two models that are not mutually exclusive have been employed to explain the mechanism of pol-II termination.

First, the 'torpedo' model suggests that cleavage at the poly(A) addition site triggers transcriptional termination by exposing the uncapped 5'-end of the downstream nascent transcript to a helicase or 5'-3' exonuclease, which would, upon catching up with the polymerase, destabilize it and cause termination (Connelly and Manley, 1988; Proudfoot, 1989). The second 'anti-terminator' model for transcription termination suggests that the polymerase has an elongation capacity that is lost upon transcription of poly(A) signals, causing the transcription complex to terminate (Logan et al., 1987). This model is consistent with experiments that show the association of 3'-end processing factors with elongating pol-II through its hyperphosphorylated carboxy terminal domain (CTD) of the large subunit. It has been proposed that dissociation of these factors from the CTD onto the 3'-end processing signals of the nascent RNA may render the polymerase complex termination competent (Dantonel et al., 1997; McCracken et al., 1997).

In the fission yeast, S. pombe, the signals required for 3'-end formation have only been defined for a couple of genes. In case of ura4 gene, it is comprised of two site-determining elements (SDEs) which position the major and minor cleavage sites and a downstream efficiency element (EE) required for efficient 3'-end formation at the upstream SDEs (Humphrey et al., 1994). The same 3'-end formation signals are required for transcription
Transcription Termination of AmA1 Genomic Clone in *S. pombe*

termination in association with a DSE which acts by pausing the polymerase (Birse et al., 1997). Termination of transcription was shown to occur 180 – 380 bp downstream of the poly(A) site, mutation in this sequence abolish 3'-end processing and affect transcription termination. Quite a reverse situation is observed in some *S. cerevisiae* genes: deletions that affect 3'-end formation do not appear to prevent efficient termination *in vivo* in *FBP1* and *in vitro* in *GAL7* (Aranda et al., 1998; Hyman and Moore, 1993). Similar to *ura4* gene, analysis of *nmt1* and *nmt2* RNA 3'-end formation signals in *S. pombe* indicates that efficient termination of transcription requires not only a poly(A) signal but also additional pause elements. Absence of such pause elements close to the poly(A) site of these genes may account for their extended nascent transcripts. Pre-mRNAs are detected to span over 4.3kb beyond the *nmt1* poly(A) site and 2.4kb beyond *nmt2* poly(A) site respectively and that is despite the presence of efficient polyadenylation signals (Hansen et al., 1998; Aranda and Proudfoot, 1999). Pol-II, in both the cases have transcribed through an entire convergent downstream transcription unit. However, the steady state levels of both the downstream genes are unaffected by the high level of *nmt1* and *nmt2* nascent transcription. The DSEs of these genes lack sequence homology, though they function in the same locations, are orientation specific and are composed of multiple, degenerate sequence elements that act together to induce full polymerase pausing. These results indicate that pausing elements could be a common feature in fission yeast genes. Further, it has been found that both poly(A) signals and pause sites are required for pol-II termination and that their relative strengths and positions determine the extent of nascent transcription. Though, *S. pombe* genome size is comparable to that of yeast, where termination occurs quite close to the poly(A) site, the fact that transcription termination in the former is produced far away from the poly(A) signal and DSE is intriguing. It is possible, therefore, that precise termination of transcription following a poly(A) signal may not be required for the expression of the many convergent genes in *S. pombe* (Oliver et al., 1992; Hansen et al., 1998).

Thus, eukaryotic pol-II termination sites remain largely uncharacterized, and no particular sequence or structural motifs have yet been identified in either lower or higher eukaryotes that can on their own direct termination of transcription by pol-II. In view of those somewhat conflicting observations about eukaryotic pol-II termination, and to understand better the *S.*
pombe termination signal, the work outlined here, intends to explain how a sequence of plant origin conforms to the requirements of a transcription terminator sequence of S. pombe and brings forward an efficient termination event. The previous study in the present investigation envisages on a prematurely terminated transcript found from expression of AmAl genomic DNA in S. pombe. The transcript of size ~450nt extends through 5'-end of the intron. As detected by northern hybridization, steady state transcript level of this small transcript is comparable with that of AmAl cDNA clone. As the poly(A) message plays a role in stability and translatability in the cytoplasm, the transcript was thought to be polyadenylated and indeed, it gave a polypeptide of size ~12kDa. Added to that presumption was the presence of higher eukaryotic poly(A) site: AAUAAA, at position 300, near the 5'-end of the intron. Hence, it was of immense interest to see whether the S. pombe 3'-end processing machinery found out matching sequences in this heterologous piece of DNA, and whether termination took place in a similar way like those of S. pombe native genes. The study assumes special significance in view of the fact that none of the three of S. pombe terminator sequences, characterized so far exhibit any homology, even though their overall functionality remains the same. Deciphering information regarding the signals perceived for this efficient (though premature!!) 3'-end formation and transcription termination of AmAl gene in S. pombe would therefore be of general interest which will help further understanding of the interplay among DNA sequence elements, RNA structures, polymerase and the factors influencing the elongation complex.

In the present investigation, it was aimed essentially to extend the observations accumulated from the transcriptional termination studies on three different S. pombe genes, to a heterologous gene (AmAl) by mapping the polymerase density across the gene sequence using a TRO assay in permeabilized yeast cells. In eukaryotes, transcription run-on assay have been employed to map sites of nascent transcription in isolated nuclei or whole cells (Connelly and Manley, 1988; Maundrell, 1990; Birse et al., 1997). TRO analysis, in which nascent RNA is radiolabeled in isolated nuclei, gives a 'snapshot' of the position of polymerases on any part of a gene (Dye and Proudfoot, 1999). In order to identify the signals required for transcription termination and pre-mRNA 3' end processing events, it was of primary interest to dissect 3' end formation signals recognized in this heterologous gene by
defining the site beyond which no further transcripts could be detected. Transcripts were further analyzed by a highly sensitive RT-PCR technique and compared the results obtained with TRO analysis. Similar procedure has been previously employed to define the region of transcription termination (Humphrey et al., 1994 and references therein). 3' end mapping of the transcript was carried out with 3'-RACE and sequencing of the amplified product and putative SDE was confirmed by a mutational study. Deletion analysis of the gene with deletion in and around the putative termination site was carried out to precisely map different elements required for 3' end processing event vis a vis transcription termination. A functional study was carried out to confirm the potency of the terminator site to bring forth a successful termination event. It was hoped that an understanding of this very event would in turn help gaining an insight into the mechanistic aspects of the concomitant event of transcription termination. A structural analysis of the terminator region was carried out to ascertain whether there is a unique structural signature present there. Finally, it was of interest to see how transcription termination of AmA1 takes place in its native system, Amaranthus. As a first step in this direction, poly(A) site in the transcript in plant was mapped.

3.2. Materials and Methods

3.2.1. Strain and primers

*S. pombe* BJ7468 (*ura4*-D18 *leu1*-32 *ade6*-M216) strain was used in the study. Cells were grown and maintained following techniques described in Moreno et al., 1991. A list of oligonucleotides used as primers is presented in the Appendix I.
3.2.2. DNA constructs

Various enzymatic deletions in the 2.55 kb AmA1 genomic DNA was carried out in pSB5.4 (Biswa, 1997). pSBDr 4.2 (a deletion from nucleotide position 535 to 1689 of the gene) and pSBDr5.1 (a deletion from nucleotide position 1321 to 1689 of the gene) clones were derived from self ligation of pSB5.4-DraI partially digested fragments of sizes ~4.2 kb and ~5.1 kb respectively (Figure 2). pSBNd4.3 (a deletion from nucleotide position 464 to 1532 of the gene) was derived from self ligation of the ~4.3 kb pSB5.4-NdeI partially digested fragment (Figure 2). The deletion clones were checked by restriction digestions and confirmed by sequencing with R1902 primer. The same deletions in pRAG were carried out by replacement cloning with BglII-NruI fragments of pSBDr4.2, pSBDr5.1 and pSBNd4.3, ligated separately to BglII-NruI digested pRAG vector backbone (Figure 3). pRABB (a deletion from nucleotide position 346 to 1675 of the gene) and pRABN (a deletion from nucleotide position 346 to 2246 of the gene) deletion clones were derived from self ligation of ~9.6 kb fragment of BglII-BclI digested pRAG and ~9.6 kb fragment of BglII-NruI digested pRAG, respectively (Figure 4). The deletion clones were checked by restriction digestions and confirmed by sequencing with R1902 primer in case of pRABN and R1044 primer in case of pRABN. In all the cases, fragments were resolved on gel, cut out respective bands, gel eluted and then ligated. The clone pΔRAG (a deletion from nucleotide position 261 to 1709 of the gene) was derived using a inverse PCR based strategy (Figure 5). PCR was set up with R261 and F1709 primers, pRAG as template using Pfu DNA polymerase (Stratagene), the amplified product of size ~10.0 kb was resolved on a 0.8% agarose gel, eluted out, ends were phosphorylated and self ligated.

pRAM275, a mutant derivative of pRAG, where the 'AA' dinucleotides at nucleotides position 275-276 of the gene was changed to 'GC', was constructed by carrying out PCR mediated site directed mutagenesis using primers F275 and R275 using template pRAG and Quickchange Site-Directed Mutagenesis Kit (Stratagene). PCR was carried out for 18 cycles each of denaturation at 95°C for 30 sec, annealing at 52°C for 1 min, polymerization at 68°C for 12 min. Similarly, clones pRABBM275 and pRABNM275 representing the above mutation in pRABB and pRABN respectively were derived using the latter two as templates.
The Rsal fragment of AmA1 intron spanning from nucleotide position 213 to 589 was cloned separately in NdeI and NruI sites of pRAC and the resultant clones were named as pRACTNd and pRACTNr respectively (Figure 9. A). pSB5.4 was digested with Rsal and the 376 bp fragment was gel eluted. Similarly, pRAC linearized with NdeI and NruI digestions were gel purified. Ends of pRAC-NdeI digested DNA were polished using DNA polymerase Klenow (NEB). Blunt end ligations were set up in 1: 6 (vector : insert) molar ratio using T4 DNA ligase (NEB) at 20°C for 16 hours.

3.2.3. Transcript 3' end mapping

Transcript 3' end was mapped using 3' RACE system (Gibco-BRL). Total RNA from pRAG transformed S. pombe cells as well as from immature seeds of Amaranthus hypochondriacus were isolated using TriPure Isolation Reagent (Boehringer Mannheim) as described in Appendix IV-4.6.

Total RNA (2.0 μg) was used for first strand cDNA synthesis carried out with Superscript II RNase H- RT (Gibco-BRL) and adapter primer (oligo-dT primer) following manufacturer's protocol. 2.0 μl of cDNA preparation was then used to seed PCR amplification carried out in 25 μl reaction volume with 0.125 U ampli-Taq DNA polymerase (Perkin Elmer), 20 pmole each of F51/F190 as forward primer and AUAP as reverse primer, 200 μM each dNTP and 2.5 mM MgCl₂. PCR amplification was carried out with initial denaturation at 94°C for 1 min followed by 30 cycles of amplification, each of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and polymerization at 72°C for 1 min, and final incubation at 72°C for 10 min. Reaction products were resolved on 1.0% agarose gel, appropriate bands were cut out from F190-AUAP PCR amplification and DNA eluted from gel piece using Qiaex II Gel Extraction Kit (QIAGEN). Ends of the PCR products were polished by DNA polymerase Klenow (NEB) and cloned at EcoRV site of pBluescript II. The clone obtained in case of 3' RACE with transcripts from pRAG transformed S. pombe was named as pRAGRACE. Similarly, pAmA1-3'R was the clone representing 3' RACE product of
transcripts from *Amaranthus*. The clones were sequenced using M13-40 Primer and Sequenase Version 2.0 DNA Sequencing Kit (USB).

### 3.2.4. Transcript analysis

The plasmid constructs described in the preceding section were introduced into *S. pombe* by the lithium acetate method (Rose *et al.*, 1990; as described in Appendix IV-4.1). Total cellular RNA was isolated using TriPure reagent (Boehringer Mannheim). 25µg of transcripts were analyzed by northern hybridization using pAmA1.3 (Raina and Datta, 1992) 1.18 kb EcoRI fragment as probe as described in Appendix IV-4.7.

RT-PCR analysis of transcripts were carried out using different primers and either of Superscript-II RT (Gibco-BRL) and ampliTaq DNA polymerase (Perkin Elmer), or rTth DNA polymerase (Perkin Elmer). In case of RT-PCR with oligo-dT primer and gene specific primer, first, cDNA synthesized was catalyzed by Superscript-II RT (Gibco-BRL) using 5µg total RNA and AP primer in a 20µl reaction volume following manufacturer's instructions. Reaction mix was treated with RNase H (Gibco-BRL). PCR amplification was carried out in 25µl reaction volume using 20 pmole each of F51 and AP primers and 2µl of the cDNA mix. Amounts of other components of the reaction added were 0.125 U ampliTaq DNA polymerase (Perkin Elmer), 200µM each dNTP and 2.5 mM MgCl₂. Amplification was carried out for 30 cycles, each of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and polymerization at 72°C for 1 min, followed by final incubation at 72°C for 10 min.

RT-PCRs with cDNA synthesized from pRAG RNA using a common forward primer F51 and different reverse primers (R261, R316, R370, R464, R680 and R1635) were carried out as described above, except in case of R680 where MgCl₂ concentration used was 3.5 mM. For reference, 0.1ng of pRAG plasmid DNA was amplified with each of the primer combinations using same conditions as was used in respective RT-PCRs. Similarly, RT-
PCRs with cDNA synthesized from pRAC RNA were carried out using F51 as forward primer and either of R1902 and R1044 as reverse primer.

RT-PCR analyses of pRAG RNA with a common forward primer F51 and different reverse primers (R261, R316, R370 and R680) were carried out using rTth DNA polymerase as described in section 1.2.5.

3.2.5. Transcription run-on (TRO) assay

Probe preparation

Different restriction fragments of 2.55 kb *AmA1* genomic DNA were used as probes for the TRO assay. These double stranded DNA probes were prepared by digesting pSB5.4 (15 μg of DNA for each digestion) with different enzymes and enzyme combinations. Restriction digestion products were resolved on 1.2% or 1.5% agarose gel depending on the size of DNA fragment(s) to be eluted. Bands specifying different fragment were cut out from the gel and DNA was eluted by phenol-freeze method. The gel piece was minced in an equal volume of phenol saturated with TE (10mM Tris.Cl, pH 8.0, 1 mM EDTA) and was frozen at -80°C for overnight. DNA was extracted out in the aqueous phase by centrifuging the mix at 10,000g for 5 min and subsequently made phenol free by re-extracting with equal volume of chloroform : isoamylalcohol (24:1). DNA was precipitated with 0.3 M NaOAc (pH 5.2) and 2.5 volumes of EtOH at -80°C, was pelleted down by spinning at 12,000g for 10 min and washed in 70% EtOH. Finally, DNA was dissolved in 50 μl of TE and gel estimated. Different *AmA1* genomic DNA fragments used as probes were described in Figure 11. A.

The fragments (30ng of each) were immobilized on a nylon filter (GeneScreen Plus, DuPont) using Hybri-Slot Manifold (Life Technologies) slot blot. DNA fragments were denatured in 0.25 N NaOH for 10 min, chilled on ice and then diluted to desired volume (50 μl) in 0.125
Transcription Termination of AmA1 Genomic Clone in *S. pombe*

N NaOH, 0.125X SSC and blotted onto membrane following manufacturer's instructions. The blot was dried at 80°C and stored at 4°C under desiccating condition.

The assay

The TRO assay was carried out following a protocol modified from those described for *S. pombe* (Maundrell, 1990; Birse *et al*., 1997). The salient steps are illustrated below.

100 ml of cultures of *S. pombe* cells (~3x10^8 cells) harbouring desirable constructs were grown in minimal medium to an OD_595 of ~1.0. Cells were harvested by centrifuging at 3000 rpm for 5 min, washed in 10 ml TMN and resuspended in 3.6 ml of H_2O. Cells were then permeabilized by treating with 400 μl of 10% sodium lauryl sulfate (SDS) on ice for 30 min. Cells were then pelleted down by centrifuging at 3000 rpm for 2 min. The pellet was recentrifuged at same rpm for another 1 min to remove residual detergent, if any.

The final pellet was resuspended in 40 μl of 2.5X transcription buffer and 60 μl of transcription buffer supplements containing dithiothreitol (DTT, 4mM), rATP, rGTP, rCTP (2mM each) and 100 μCi of (α-^{32}P) rUTP (3000Ci/mmol, BARC) and transcription was allowed to proceed for 5 min at 25°C. Cells were collected by centrifugation at 3000 rpm for 2' at 4°C, washed briefly with TMN and finally proceeded for total RNA isolation using Tripure Isolation Reagent (Boehringer Mannheim) (Appendix IV-4.6).

Total RNA extracted was partially hydrolyzed by treating with NaOH (0.2 M) for 5 min at 4°C and neutralized with Tris-HCl, pH 7.2 (0.2 M). Samples were then hybridized to the immobilized probes following the protocol used for northern hybridization (Appendix IV-4.7). Filters with immobilized probes were washed in 2X SSC for 1 min and prehybridized in 10 ml of 5X SSC, 50% (w/v) deionized formamide, 10% (w/v) dextran sulphate - Na salt (MW 500,000), 1% (w/v) SDS at 42°C for 4 hrs. Partially hydrolyzed RNA transcripts were added to the prehybridization solution along with denatured sheared salmon sperm DNA at a concentration 200 μg/ml. Hybridization was carried out for 24 hrs at 42°C. Filters after
hybridization were rinsed with 0.2X SSC, 0.1% SDS followed by washing in same solution twice at 42°C for 20 min each, and signals were autoradiographed using Kodak (X-OMAT) AR films. Signals were quantified in Gel Doc 1000 (Bio-Rad).

Solutions

20X SSC: 3.0 M NaCl, 0.3 M Sodium citrate dihydrate
TMN: 10 mM Tris-HCl pH 7.4, 5 mM MgCl₂, 100mM NaCl
2.5X Transcription buffer: 50 mM Tris-HCl pH 7.7, 500 mM KCl, 80 mM MgCl₂

3.2.6. In vitro transcription interference assay

In vitro transcription reactions were set up in 50 µl reaction volume using pSBNd4.3 linearized with Ndel as template DNA and T3 RNA polymerase (Gibco-BRL). The reactions were set up essentially as described in the manufacturer’s protocol (Gibco-BRL) with 100 ng of template DNA, 5 mM DTT, 1 mM rATP, 0.5 mM each of rCTP and rGTP, 1 µl of RNase inhibitor, 2.0 µl (α-32P)UTP (Amersham) (20µCi), 1.0 µl of T3 RNA polymerase and 10.0 µl of 5X Transcription buffer (0.2 M Tris-HCl, pH 8.0, 40 mM MgCl₂, 10 mM spermidine-(HCl)₃, 125 mM NaCl) at 37°C for 30 min. Further, the reaction had 20 pmole of either of the following antisense oligonucleotides – R261, R316, R370 and R464. Similarly, reactions were set up varying the amount of R316 from 1 pmole to 40 pmole. 5.0 µl of the reactions were analyzed on 5% polyacrylamide(29:1)-8M urea gel. As reference molecular weight marker, 100 bp DNA ladder (NEB) was end labeled using DNA polymerase klenow and (α-32P)dATP and loaded along with the samples. After the run, the gel was treated with gel fixing solution (10% acetic acid, 10% MeOH) for 10 min, lifted onto a Whatman 3 mm filter paper and dried. The gel was exposed to Kodak X-Omat K (XK5) film.
3.3. Results

3.3.1. Mapping of Poly(A) site of the transcript of AmA1 genomic clone (pRAG) in S. pombe

To map the poly(A) site of the prematurely terminated transcript, 3'-RACE was carried out. Total RNA was isolated from the cells transformed with pRAG and subjected to RT-PCR using a gene specific primer and an oligo(dT) primer. The major DNA product was then subcloned and sequenced. As presented in Figure 1, sequencing of the 3'-RACE product revealed a sequence TA starting at position 299 of the gene as poly(A) site. The sequence conforms to the most favourable cleavage site in S. pombe, usually characterized by the sequence Y(A)n. An inspection of the sequence in AmA1 gene surrounding this site revealed the presence of an eukaryotic positioning element AATAAA generally positioned 10-30 nt upstream of the poly(A) site (Guo and Sherman, 1996), at a position 24 nt upstream. This sequence element with the similar positioning with respect to the cleavage site in the eukaryotic consensus, suggests it as a candidate SDE in S. pombe which may involve in defining the cleavage site.

3.3.2. Mapping of sequences required for transcription termination and mRNA 3' end processing in S. pombe

Northern blot analysis of total RNA extracted from S. pombe transformants carrying different deletions in the AmA1 genomic DNA using 1.18 kb AmA1 cDNA as probe indicated a varying effect of different intronic deletions on transcriptional end product (Figure 6. B). Results revealed that the clone pΔRAG, has a deletion in the region of 261-1709 (Figure 5), abolished the premature termination event occurred in case of the parental clone. The clone
Figure 1. Poly(A) site mapping of the premature transcript using 3' RACE strategy: A. Electrophoretic separation of 3' RACE products. RT reaction was carried out using 3' RACE kit (GIBCO BRL) with total RNA isolated from *S. pombe* transformed with pRAG. 2 µl of RT reaction mix was used to seed PCR amplification in 25 µl reaction volume carried out for 30 cycles using ampli-Taq DNA polymerase (Perkin Elmer) and F51/F190 and AUAP primers. Products were analyzed on a 1.2% agarose gel. M: 1 kb DNA ladder (GIBCO BRL). B. Cloning of the 3' RACE product in pBluescript II. F190-AUAP 3' RACE product was gel purified using Qiaex II Gel Extraction Kit (QIAGEN), end polished and cloned in the EcoRV site of pBluescript II. The resultant clone was named as pRAGRACE C. Sequencing of the 3' RACE product. pRAGRACE clone was sequenced using M13 -40 Primer and Sequenase Version 2.0 DNA Sequencing Kit (USB). Sequence read is shown in the right panel.
pRABB, a deletion in the region 346-1675 (Figure 4), yielded two different kinds of transcripts, one prematurely terminated product of size ~450 nt matching with transcripts from pRAG clone and the other one of size comparable to the size of the deletion clone coming as a read-through transcript. The clone pRABN, a deletion in the region of 346-2246 (Figure 4), indicated three different transcripts on the northern blot. The longest one was a transcript of size comparable to the deletion clone and the shortest one being of the same size with the transcript from the parental clone; while the transcript of intermediate size could probably be a product coming out of an aberrant splicing event took place in the read-through population. All other deletion clones, namely, pRANd (464-1532 deletion, Figure 3), pRADr10.0 (535-1689 deletion, Figure 3) and pRADr10.9 (1321-1689 deletion, Figure 3) yielded single transcript of size alike to the transcript from pRAG. So, except the clone pΔRAG, all the other deletion clones invariably gave only a prematurely terminated product. Only exception was in case of the clones pRABB and pRABN, where both exclusively produced two kinds of transcripts, truncated transcript and a read-through transcript. Analysis of these results points out two regions which may play significant role in the termination event, one around position 346 and the other around position 464. pΔRAG, a deletion from position 261 onwards completely abolished the premature termination event that occurred in a downstream region of an undeleted clone. pRABB and pRABN (deletions from position 346 onwards) clones yielded both kinds of transcripts indicating that the premature termination was not a full proof one, and hence, the sequence flanking position 346 must play a role in the termination event, probably by enhancing its efficiency. The sequence relevance of the region between 346 to 464 draws more attention as in the clones where this region was retained, e.g. pRANd, pRADr10.0 and pRADr10.9, the premature termination occurred up to completion producing only the small transcript. These findings are reminiscent of the efficiency element and the role played by DSEs found in S. pombe terminator sequences (Birse et al., 1997).

A fine look on the northern blot reveals more about the transcripts. There was an additional small transcript of size ~600 nt, though less in amount, in case of pRAG, pRADr10.0, pRADr10.9 clones and very less in case of pRANd clone. No such transcript was there in case of pRABB and pRABN. This additional transcript might have come from the clone
Figure 2. Deletions in the intronic region of AmA1 genomic DNA. pSB5.4, the 2.55 kb AmA1 genomic clone at EcoRV site of pBluescript II, was partially digested with Dral and NdeI separately and suitable fragments eluted from the gel and self ligated. Resultant clones were named as pSBDr4.2, pSBDr5.1 and pSBNd4.3, respectively. Scheme presented is not drawn to scale. M: 1 kb DNA ladder (NEB).
Figure 3. Deletions in the intronic region of AmAl genomic clone in pRAG. pSBDr4.2, pSBDr5.1 and pSBNd4.3 clones were digested with BglII and NruI and the fragments of sizes ~750 bp, ~1.5 kb and ~850 bp were gel purified and ligated separately to ~9.4 kb vector backbone prepared from pRAG digested with same enzyme combination. Resultant clones were named as pRADr10.0, pRADr10.9 and pRANd respectively. Scheme presented is not drawn to scale.
Figure 4. Deletions of *AmAl* genomic clone in pRAG. A. pRAG was digested with BgIII and BclI and the 10.0 kb fragment containing vector backbone was gel purified and self ligated. The resultant clone was named as pRABB, deletion in the region spanning from position 346 to 1675 of the gene. B. pRAG was digested with BgIII and NruI and the 9.6 kb fragment containing vector backbone was gel purified, blunt ended by Mung bean nuclease treatment (NEB) and self ligated. The resultant clone was named as pRABN, deletion in the region spanning from position 346 to 2246 of the gene. Schemes presented is not drawn to scale.
Figure 5. Deletion of the intronic region spanning from position 261 to 1709 of the \textit{Am4I} genomic DNA using a restriction free PCR based deletion technique. 

A. Schematic representation of the technique used. B. PCR amplified deletion product in the ethidium bromide stained gel. Lane 1. Amplified product, lane 2. 1 kb ladder (Gibco-BRL), lane 3. Single cut template plasmid pRAG. PCR was performed in a 25.5 µl volume using 5 ng of pRAG as template DNA. A reaction mixture was prepared containing 1X Pfu polymerase buffer (Stratagene), 62.5 ng of each of forward primer F261 and reverse primer R1709, 200 µM each dNTPs and 1.25 U Pfu DNA polymerase (Stratagene). Cycle parameters were initial melting at 95°C for 30 s, followed by 16 cycles each consists of 95°C for 30 s, 50°C for 1 min, 68°C for 23 min. The reaction was subsequently electrophoresed on an 0.7% agarose gel (B) and amplified product of size ~10.0 kb was gel purified using Qiagen II gel purification kit (Qiagen). 5' end of the product was phosphorylated using Polynucleotide kinase (NEB) and self ligated using T4 DNA ligase (NEB). C. Confirmation of deleted clone by restriction analysis. Lane 1. Uncut pRAG, lane 2. Uncut pΔRAG, lane 3. 1 kb ladder (Gibco-BRL), lane 4. Nrul digested pRAG, lane 5. Nrul digested p ΔRAG. N.B., primers were designed such that after the ligation of the inverse PCR product, it should create a Nrul site, in addition to the existing one.
Figure 6. Analysis of deletions in the intronic region of AmAl genomic clone in S. pombe. A. Representation of the deletions in the 2.55 kb AmAl genomic clone (pRAG) used in the mapping of sequence elements involved in transcription termination. B. Northern hybridization of total RNA isolated from S. pombe transformed with constructs illustrated above. 25 μg RNA was separated on a 1.2 % denaturing formaldehyde-agarose gel, transferred onto nylon membrane and probed with 1.18 kb AmAl cDNA. Positions of RNA molecular weight markers (New England Biolabs) have been shown in the right track.
Transcription Termination of *AmA1* Genomic Clone in *S. pombe*

carrying the region of 346-464. These small transcripts, polyadenylated at different poly(A) sites, may have a common or have two entirely different SDEs, needs to be worked out.

3.3.3. RT-PCR analysis of transcript polyadenylation

An RT-PCR analysis was carried out with total RNA preparation from *S. pombe* cells transformed with different constructs using an oligo-dT primer as reverse primer and F51 as forward primer (Figure 7. A). The results obtained clearly revealed the presence of three different polyadenylated species of transcripts in case of pRAG. Among them, a product of size ~340nt was the major product, while two other products of sizes ~420nt and ~440nt respectively were minor products. These bands could be correlated with the finding in case of northern hybridization which clearly indicated two different small transcripts of sizes ~450nt and ~600nt respectively. Probably the transcripts corresponding to the two low yield RT-PCR products could not be resolved in northern, and hence, merged to form the ~600nt band. pRABB yielded only single amplification product of size ~340nt, common to pRAG. The amplification of the read-through transcript present in pRABB, could not be met probably due to the conditions used in the amplification. In case of pRABN, there were three distinct polyadenylated products tallying with the respective sizes of the transcripts estimated earlier from northern blot analysis. The major truncated transcript was of ~340nt while ~750nt represented the read-through transcript and the ~600nt transcript was its aberrant spliced product. pRAC (*the AmA1* cDNA clone), yielded ~1.0 kb amplified product which could be correlated well with the equivalent transcript of size ~1.1kb in the northern hybridization.

The results of this RT-PCR analysis conformed well with the results obtained from the northern hybridization and unequivocally demonstrated that the different transcripts were, indeed, polyadenylated.
Figure 7. A. RT-PCR analysis of transcripts using oligo-dT antisense primer. RT reaction was carried out using Superscript- II RT (3' RACE kit, GIBCO BRL) with total RNA isolated from *S. pombe* transformed with pRAG, pRAC, pRABB and pRABN, respectively. 2 μl of RT reaction mix was used to seed PCR amplification in a 25 μl reaction volume. PCR reaction was carried out for 30 cycles using ampli-Taq DNA polymerase (Perkin Elmer) and F51 and AP primers; products were analysed on a 1.2% agarose gel. M: either 1 kb DNA ladder (GIBCO BRL) or 100bp DNA ladder (NEB). Arrows indicate different polyadenylated species of RNA. B. Sequence of the *AmAl* intronic region acted as transcription terminator in *S. pombe*. 213-589 Rsal fragment (size: 376nt) was used to carry out a functional assay to show effectiveness of the terminator sequence. Figure shows positions of different reverse primers used in fine mapping and structural analysis of the terminator.
3.3.4. Mapping of site determining element (SDE)

Mapping of poly(A) site of the truncated transcript of *AmA1* gene in *S. pombe* indicated the sequence AATAAA at position 273 to be a putative SDE. The sequence was mutated to AAGCAA and to study its effect on transcription, pRAG and the deletion clones, pRABB, pRABN and pRANd, all were mutated in the putative site and named as pRAM275, pRABBM275, pRABNM275, pRANdM275 respectively and introduced in *S. pombe*. Total RNA isolated from the transformants were analyzed on northern hybridization using 1.2 kb *AmA1* full length cDNA as probe (Figure 8). Results revealed that the mutation significantly affected the yield of the small transcript of size ~450 nt. The clones where two small transcripts of sizes ~450 nt and ~600 nt respectively were produced with the former one predominant (i.e., pRAG, pRANd and pRABN), mutation altered the ratio in favour of the less predominant one (Figure 8). This clearly indicates that the sequence AATAAA acts as SDE in *S. pombe*, though with a varied efficiency as the mutation could not abolish polyadenylation completely at the downstream poly(A) site.

3.3.5. Functional assay of the transcription terminator mapped

To study the functionality of the sequence in the *AmA1* gene which houses the putative elements required for polyadenylation and transcription termination in *S. pombe*, sequence spanning from position 213 to 589 (named as Ter, henceforth and represented in Figure 7. B) was cloned at two different positions in the *AmA1* cDNA and the effect was assayed. Both the clones namely pRACTNd (Ter clone in the NdeI site at position 492 of *AmA1*, Figure 9. A) and pRACTNr (Ter clone in the NruI site at position 714 of *AmA1*, Figure 9. A) when introduced in *S. pombe* could bring successful premature termination events in case of the cDNA clone. This was revealed by the northern hybridization carried out with total RNA isolated from the transformants using 1.18 kb *AmA1* cDNA as probe (Figure 9. B). Drifting of the transcript size in both the clones to ~650 nt and ~900 nt respectively from the size
Figure 8. Analysis of putative SDE mutations of Am41 genomic clone and the deletion clones in S. pombe. A. Northern hybridization of total RNA isolated from S. pombe transformed with different mutant clones. 25 µg RNA was separated on a 1.2 % denaturing formaldehyde-agarose gel, transferred onto nylon membrane and probed with 1.18 kb Am41 cDNA. Lower panel is the impression of the gel stained with ethidium bromide prior to the transfer. Positions of RNA molecular weight markers (New England Biolabs) have been shown in the right track. B. Bar graph of the data obtained in the northern hybridization. The signals were quantified in a documentation system (Gel Doc 1000, BioRad). The absolute values were corrected for background hybridization and plotted against clones, with different colours indicating different size of the transcripts.
Figure 9. Functional analysis of AmAl intronic sequence recognized as transcription terminator. **A.** Cloning of putative terminator sequence. The 376 kb Rsal fragment (213-589), was ligated to Ndel cut and Klenow filled pRAC. The resultant clone was named as pRACTNd. Separately, the same fragment was ligated to NruI cut pRAC and the clone obtained was named as pRACTNr. **B.** Northern hybridization of total RNA isolated from *S. pombe* transformed with different clones. 25 μg RNA was separated on a 1.2 % denaturing formaldehyde- agarose gel, transferred onto nylon membrane and probed with 1.18 kb AmAl cDNA.
~1500 nt, proves the 'ter' functionality and thereby, demonstrates the efficacy of the terminator sequence mapped.

3.3.6. Detection of transcripts downstream of the poly(A) site

mRNA 3' end formation in *S. cerevisiae* has been proposed to result from coupled transcriptional termination and pre-mRNA processing events. With a belief that the paradigm fits in case of the termination event which the present investigation is studying, an analysis of the RNA 3' end formation signals recognized by *S. pombe* in the *AmA1* gene was carried out by defining the site at which no further transcripts could be detected in the 3' flanking region of the poly(A) site. To obtain this information, the transcripts were analyzed by a sensitive RT-PCR technique (Figure 10). DNA primers were hybridized to the transcripts from pRAG cells and the reverse transcribed cDNA products were then amplified using PCR. Antisense oligonucleotides which were picked up from different regions spanning the whole intron are illustrated in Figure 10. A, and sequences presented in Appendix I. These primers were used for reverse transcription of DNase treated total RNA which was isolated from *S. pombe* transformed with pRAG. These reverse transcripts were PCR amplified using the same antisense oligonucleotide and a common 5' oligonucleotide which had been radioactively labeled. A measure of the relative amount of transcripts were made by comparing the counts incorporated into the PCR products separated electrophoretically (Figure 10. B and C). To avoid any erroneous estimation due to varied RT-PCR efficiency for each pair of primers, a known amount of plasmid DNA was PCR amplified using same amount of the primer combination as a reference and the amount of transcripts were expressed in percentage amplification relative to the reference (set at 100). The amount of DNA used in the reference reaction was selected such that all the reverse primers used in the study gave near equal amplification in terms of number of counts. Hence, the relative amount of the transcript expressed in percentage after normalization should correlate with the amount of transcripts extended up to the antisense oligonucleotide binding site.
Figure 10. Detection of transcripts downstream of poly(A). A. A representation of the relative positions of each of the antisense oligonucleotides used for RT-PCR analysis with respect to the 2.55 kb *AmA1* genomic DNA. B. Electrophoretic separation of RT-PCR products detected across downstream region of the poly(A) site. RT reaction was carried out using Superscript-II RT (3' RACE kit, GIBCO BRL) with total RNA isolated from *S. pombe* transformed with pRAG. 2 μl of RT reaction mix was used to seed PCR amplification in a 25 μl reaction volume, reaction carried out for 25 cycles using ampli-Taq DNA polymerase (Perkin Elmer) and a common forward primer, F51 (5' end labeled) and different antisense primers. As reference for semi-quantification of differential amplification, 0.1 ng of pRAG DNA was similarly PCR amplified. 10 μl products were analyzed on a 1.5 % agarose gel. IVT-RNA: *In vitro* transcribed RNA. M: 1 kb DNA ladder (GIBCO BRL). C. Bar graph of the data obtained in B. The bands were excised and radioactivity (Cerenkov count) measured, plotted as the percentage of the relative count obtained in case of plasmid DNA as template (set at 100%).
From such RT-PCR analysis, transcripts were detected past the previously mapped poly(A) site, although at a much reduced level compared with the level of transcripts detected near downstream of the poly(A) site. Amplified product was detected in the present experimental set up, up to the reverse primer R680, whose 5' end starts at position 680. Analysis of this finding in light of what had been found in the northern blot, where two transcripts of size ~450 nt and ~600 nt respectively were end products, entails that transcription termination must have occurred close to the most distal RNA 3' end formation element roughly defined in that study. These observations may implicate signals within this region within position 680 as having a role in transcription termination.

A similar study in case of the expression of AmA1 cDNA clone in S. pombe, relative amounts of transcripts at two different primer binding positions were estimated (Figure 10. B). The analysis indicates a fall in full length transcript that could be amplified by the primer R1044 positioned at the end of cDNA. Relatively less amplification by R1044 could be interpreted as due to less primer binding in the RT step as the control reaction with in vitro transcribed RNA as template showed near equal amplification with both the primers R1902 (located far upstream to R1044) and R1044. Polyadenylation of the pRAC transcripts taken place at a position in the R1044 primer binding site, seems to be a good proposition for the moment. This is on line with the inputs from the northern analysis which clearly indicated a transcript size ~1.1 kb from the clone (Figure 6).

3.3.7. Detection of extended nascent transcription using TRO assay

In higher eukaryotes, transcription run-on (TRO) assays have been employed to map sites of nascent transcription in isolated nuclei or whole cells. A TRO assay was employed to map polymerase density across the 3'-flanking region of the mapped poly(A) site for the premature termination event. Briefly, detergent-permeabilized yeast cells are incubated in a moderately high salt transcription buffer containing (α-32P)UTP for a short period. The conditions permit transcriptionally engaged polymerases to elongate a short distance and incorporate radioactive label (Birse et al., 1997, Hansen et al., 1998). Total RNA is then
Figure 11. Transcription run-on (TRO) analysis of 2.55 kb genomic clone of \textit{AmAl} in \textit{S. pombe}. A. Schematic representation of different overlapping fragments of the 2.55 kb \textit{AmAl} genomic DNA used as probes in the assay. Restriction enzymes used to prepare the fragments are shown along with their respective positions. B. TRO blot. 30 ng of each of the probes were blotted and hybridized with nascent transcripts isolated from pRAC and pRAG transformed \textit{S. pombe} cells as described in the materials and methods. The signals obtained after the high-stringency wash are shown.
Figure 12. Graphical display of the TRO profile. A. Sketch of 2.55 kb genomic DNA of AmA1 in S. pombe expression cassette. Positions (start points) of each TRO probes (represented by numbers) used in the study are depicted in the lower panel. Poly(A) sites and putative SDE in the intronic region of AmA1 genomic DNA that is recognized in S. pombe, are shown. B. The signals of the TRO blot are displayed graphically after quantitation. The graph is drawn to scale such that the width of each bar reflects the length of the probe and the height indicates respective signal across the probe. I and II are the plots of TRO signals obtained from pRAC and pRAG transformed S. pombe cells respectively. III displays signals across the probes in case of pRAG transformed cells which are corrected and expressed relative to probe-1.
isolated, partially hydrolyzed and hybridized to immobilized denatured DNA probes to localize the active polymerase complexes. The signal for each probe is proportional to the average polymerase density across the DNA fragment, within the yeast population. If the pulse-labeled RNA is hybridized to contiguous single-stranded DNA probes spanning the 3'-flanking region of the poly(A) site in question, a profile of polymerase density could be obtained.

TRO analysis was performed on the pRAG and pRAC transformed S. pombe cells, both, where the latter served as control (Figure 11. B). The TRO profile obtained is presented graphically in Figure 12. The probes used spun the whole 2.55 kb of the AmA1 genomic clone and were mostly overlapped with one another (Figure 11. A). TRO analysis indicated that a high level of active polymerases accumulated over probes- 4, 6, 8 and 9 of the genomic clone, showing strongest signal with probe-9. Delineation of the probes to the DNA sequences they cover, points out two distinct regions, one spanning from 338 to 590 and the other from 589 to 949. The signals represent the elevated level of polymerase density over these two regions where elongating polymerases probably accumulate in response to transcriptional pause signals in a way that is characteristic of DSE. The signals obtained in case of pRAC transformed cells revealed highest polymerase density with probe-12, i.e., the probe which spans the whole exon-2 of the gene, probably either due to the larger probe size or due to elevated rates of polymerase elongation across the DNA, or due to both. In that region, there should not be any DSE like element present as a previous experiment (Figure 10) roughly assigned a sequence around R1044 binding site (position around 2545 with respect to the genomic clone and position 1011 with respect to the cDNA clone) as a putative poly(A) site in case of pRAC.

3.3.8. RNA secondary structure analysis

RT-PCR analysis of transcript using reverse primers: R261, R316 and R370 (named with respect to the position of the 5' end), separately with a common forward primer F51, indicated maximum amplification with R370, in comparison to those others with R261
showing minimum (Figure 13. B). It was quite intriguing keeping in view their comparable melting temperatures. The variable pattern of amplification was found to be same regardless the protocol used in the RT-PCR and hence, could be interpreted as an intrinsic property of the transcript. This finding motivated the study towards secondary structural analysis of the RNA in that region. Moreover, other experiments designed to address transcription termination and 3' end processing, assigned significance to this very region as it housed SDE and poly(A) site, besides stipulating the presence of DSEs like element that stalls elongating polymerase.

A structural analysis of the RNA sequence in the region between nucleotide 213 and 589 using the program 'RNA fold' indicated three stem loop structures arranged in tandem in region between 245 to 367 (Figure 13. A). To substantiate the presence of secondary structure in the region which probably, plays a crucial role in 3' end processing, an assay was carried out which was subsequently named as 'in vitro transcription competition assay', where release of RNA from transcribing complexes upon challenging with oligonucleotides was measured. When a transcribing bacterial RNA polymerase (T3/T7 or SP6 RNA polymerase) encounters secondary structure in the transcribed RNA, transcription termination takes place and RNA polymerase falls off releasing the transcript. If a secondary structure formation is counteracted by an oligonucleotide complementary to RNA sequence, RNA polymerase should produce more of 'run off' transcripts. Precisely that was the working principle of the assay. Results obtained from the experiment indicated a maximum level of read through transcript in case of challenge with R316 in a concentration dependent manner (Figure 13. C). The experiment clearly indicated the presence of secondary structure in the region as was predicted by RNA fold analysis. Role of secondary structure in this very region of the transcript which probably falls in the realm of the first DSE, in the transcription termination event needs to be worked out in detail.
Figure 13. Structural analysis of the transcript. A. Secondary structure prediction of the transcript flanking poly(A) site. Sequence spanning nucleotide position from 1 to 464 of the AmA1 2.55 kb genomic DNA was analyzed using RNA Fold program. Only structures of stem-loops predicted in the region are shown. B. RT-PCR analysis with different reverse primer in the stem-loop region. Electrophoretic separation of RT-PCR products generated after 20, 25, 30 PCR amplification cycles from total RNA isolated from S. pombe transformed with pRAG. RT-PCR reactions were carried out in 25 μl reaction volume with 1.0 μg of total RNA using rTth DNA polymerase RT-PCR kit (Perkin Elmer) using 5’ end labeled F51 and one of the four different reverse primers R261, R316, R370, R680 and 10 μl of the products were analyzed on a 1.5 % agarose gel. M: 1 kb DNA ladder (GIBCO BRL). Amplification products from 25 cycle amplification were excised and radioactivity (Cerenkov count) measured, represented in the graph shown in the side panel. C. In vitro transcription termination competition assay. In vitro run off transcription using T3 RNA polymerase (GIBCO BRL) from the template pSBNd4.3 linearized with Ndel was carried out in presence of different reverse primers (Primers used and their amounts are indicated above the tracks). Reactions were carried out in a 50 μl reaction volume using α-(32P)UTP and 5.0 μl of reactions were analyzed on a denaturing polyacrylamide gel. Relative positions of the antisense primers used in the competition assay with respect to the RNA secondary structure is shown in A.
3.3.9. Poly(A) site mapping of *AmAl* in its native system

It was of interest to see how transcription termination of *AmAl* takes place in its native system, *Amaranthus*. To add to this direction, primarily, it was decided to map the poly(A) site in the transcript in plant. Total RNA was isolated from immature seeds of amaranth and was subjected to 3' RACE. The product was subsequently cloned and sequenced (Figure 14). Reading of the sequence unequivocally indicates a sequence 'TA' at position 159 nt downstream of the translational stop codon (positioned at 2448 of the 2.55 kb genomic clone pSB5.4, keeping ATG at position 1). Sequence analysis of the 3' UTR of the transcript identifies an eukaryotic polyadenylation signal consensus AATAAA, at a position 98 nt upstream of the mapped poly(A) site. This sequence probably is the candidate positioning sequence element in case of the *AmAl* gene.

3.4. Discussion

The aim of the study was to work out the finer details that shrouds the event of premature termination in the *AmAl* gene in *S. pombe*. Though RNA polymerase II transcriptional termination is a poorly understood process, it has been known for some years that the presence of a poly(A) signal at the 3' end of a gene plays a role both in determining the site of termination as well as mRNA 3' end processing (Proudfoot, 1989). However, the molecular connections between termination and polyadenylation are becoming clearer only recently and project RNA processing as a co-transcriptional event. Two main sequence elements are likely to be involved in Pol II termination, an upstream polyadenylation signal and a downstream element involved in polymerase pausing (Aranda and Proudfoot, 1999). The present study is intended to analyze the transcription termination event from the dual perspectives of 3' end processing and polymerase dissociation, drawing parallels with the basic paradigm.
Figure 14. Poly(A) site mapping of the *AmAl* transcript in plant using 3' RACE strategy. A. Electrophoretic separation of 3' RACE products. RT reaction was carried out using 3' RACE kit (GIBCO BRL) with total RNA isolated from immature seeds of *Amaranthus hypochondriacus*. 2 μl of RT reaction mix was used to seed PCR amplification in 25 μl reaction volume carried out for 30 cycles using ampli-Taq DNA polymerase (Perkin Elmer) and F51/F190 and AUAP primers. Products were analyzed on a 1.2% agarose gel. M: 1 kb DNA ladder (GIBCO BRL). B. Cloning of the 3’ RACE product in pBluescript II. F190-AUAP 3’ RACE product was gel purified using Qiaex II Gel Extraction Kit (QIAGEN), end polished and cloned in the EcoRV site of pBluescript II. The resultant clone was named as pAmAl-3’R. C. Sequencing of the 3’ RACE product: pAmAl-3’R clone was sequenced using M13-40 Primer using Sequenase Version 2.0 DNA Sequencing Kit (USB). Sequence read is shown in the right panel.
Towards this end, firstly the poly (A) site of the transcript was mapped using 3' RACE strategy and subsequent sequencing of the 3' RACE product which revealed 5'-TA-3' dinucleotide at position 299-300 as the poly(A) site. The poly(A) site had the invariable 5'-Y(A)n-3' configuration which is often located between 10-30 nucleotides downstream of polyadenylation signal in eukaryotes. Deletion analysis in the intronic region downstream of the poly(A) site demonstrated a varied role of different sequences over the formation of the prematurely terminated transcript. Further, a detail analysis of the results demonstrated that the genomic clone yielded two smaller transcripts of sizes ~450 nt and ~600 nt respectively. Different deletions had a diverse effect on the formation of these two products. It was quite clear that deletion from position 261 onwards to far downstream in the intron resulted in total read-through transcription, abolishing the termination event encountered by the undeleted clone. Deletion from 346 to the downstream intronic region resulted not only read-through transcription but also premature termination, but yielded only one small product of size ~450 out of the latter event. So, there must lie a sequence flanking position 346 which is important for the premature termination event as well as for the generation of the ~600 nt product and deletion of which reduces the efficiency of premature termination giving read-through transcripts. The deletion from position 464 onwards into the intronic region could give complete premature termination but with very less yield of the ~600 nt small transcript. It was quite clear that though the region 261-464 was sufficient for the premature termination, it was not sufficient for the generation of the second polyadenylated small transcript. The deletion from position 535 onwards, within the intron could bring about a complete premature termination event entailing the region 261-535 which houses a complete repertoire of cis elements required for the event and generation of the two small products.

Subsequently, it emerged from an RT-PCR study on polyadenylation that there were three different polyadenylated species of transcripts in case of the AmA1 genomic clone in S. pombe. Among them, a product of size ~340 nt was the major product, while two other products of sizes ~420 nt and ~440 nt respectively were minor products. The result could be correlated with the finding in northern hybridization which clearly indicated two different small transcripts of sizes ~450 nt and ~600 nt respectively in case of AmA1 genomic clone in S. pombe, probably the transcripts corresponding to the two low yield RT-PCR products.
Transcription Termination of *AmA1* Genomic Clone in *S. pombe*

merged to form the ~600nt band. There might be different poly(A) sites involved in the species diversity of the terminated transcripts. It seems equally probable that different polyadenylation signals within the foreign gene intronic region, have been recognized in the heterologous piece of DNA by *S. pombe* system. Mutational analysis effectively entailed that the eukaryotic polyadenylation signal consensus AATAAA, present 25nt upstream of the poly(A) site worked in *S. pombe* as its equivalent site determining element (SDE). A mutation in the sequence could not abolish completely the formation of prematurely terminated transcript, demonstrating that the sequence was not indispensable. Mutation in the putative sequence altered the relative amount of different transcripts. It can be argued that there may be more than one polyadenylation signals clustered around this sequence which gets more activated in the dysfunction of the first signal. If that is the case, the situation observed in this instance will appear to be similar to that in *S. cerevisiae* where multiple positioning elements are found directing cleavage at several clustered sites 15-30 nt downstream (Heidmann *et al.*, 1992; Guo and Sherman, 1995, 1996). Previous studies on *nmt1*, *nmt2* genes of *S. pombe* revealed the presence of an A-rich sequence located ~15 nt upstream of the poly(A) site and comparison of their sequences with the proposed consensus positioning element (AAAAAAA or AAUAAA) in *S. cerevisiae* and the similar positioning with respect to the cleavage site suggests an A-rich element as a candidate SDE in *S. pombe* which may act to define the cleavage site (Hansen *et al.*, 1998). Precisely, that was the scene in case of the expression of the genomic clone and the *S. pombe* machinery found out the matching counterpart(s).

Keeping in view that the sequence in the region spanning 261-535 was, *per se*, sufficient for the premature termination event, a functional analysis of this region was planned to assess the efficacy of this region in pursuing similar event in different genetic environment. To do the same, the RsaI digested 376 bp fragment spanning the region from positions 213 to 589 was cloned in two different positions of the *AmA1* cDNA clone and introduced in *S. pombe*. The results unequivocally demonstrated the effectiveness of the region, bringing successful termination events in both the clones. The study further reports the cloning of sequence that could be used as a potent transcription terminator in *S. pombe*.
The results accumulated from the array of analyses bring relevance to two different regions that could effectively bring forth the termination event, deletion of which resulted in accumulation of read through transcripts. Deletion of the region spanning 346-464 could be implicated in the reduction of the efficiency of the termination giving read-through transcription but sufficient for the major polyadenylation event; the region 346-535, per se, seemed to be sufficient for the termination event and for the formation of both major and minor polyadenylation products, and the region 464-535 probably had a role in the generation of the low yield polyadenylated transcript over and above the termination process.

The SDE and the poly(A) site which were involved in the major polyadenylation events were dissected out and found to be sequences located at positions 273-278 and 299-300 respectively. The equivalent cis elements involved in bringing the minor polyadenylation event in all probability, might fall in the region spanning 346-535 and more precisely, in the region flanking position 464. Probably, the SDE of this event falls in the region near upstream of position 464. These findings seems to be analogous to those cis elements defined for the ura4 gene in S. pombe and involved in 3' end formation. The signals comprise of two SDEs which position the major and minor cleavage sites and a downstream efficiency element (EE) required for efficient 3' end formation at the upstream SDEs (Humphrey et al., 1994). Subsequently, it has been shown that production of mRNA 3' termini in S. pombe involves at least two functionally distinct elements: SDE and EE, working in concert, where the EE is replaceable with a second copy of SDE without altering its functionality.

The findings raise some other interesting questions as well. First, what sequences comprise the site determining element? Computational analysis of the 3' ends of S. pombe genes has not revealed any sequence homologies (Humphrey et al., 1994 and references therein). It sounds interesting to establish the evolutionary relationship of the bipartite poly(A) signal observed in S. pombe to lower and higher eukaryotes. It has been demonstrated that S. cerevisiae Poly(A) signals function in S. pombe while mammalian poly(A) signals do not. However, the bipartite arrangement of the SDEs and EEs, a feature common in S. pombe, is very reminiscent of mammalian poly(A) signals where two sequence elements, the hexanucleotide AAUAAA and a U- or GU-rich downstream signal are necessary for efficient 3' end formation (Humphrey et al., 1994; Wahle and Keller, 1992; Guo and Sherman, 1995).
Modular architecture of plant poly(A) signal, which comprises of near upstream element (NUE) and far upstream element (FUE) closely resemble that of yeast (Mogen et al., 1990, 1992). Components of animal, *S. cerevisiae*, plant and *S. pombe* poly(A) signals are presented in Figure 15 and a model has been proposed to show the functionality of the *AmA1* intronic region in 3' end processing.

Fortified with this array of information regarding 3' end processing, the study made an effort to understand the other relevant events happening almost co-laterally to this event. Few other studies in *S. pombe* system emphasized the importance of both poly(A) signals and transcriptional pause sites for efficient termination (Aranda and Proudfoot, 1999). These studies further revealed that *S. pombe* polyadenylation signals are often degenerate and redundant but they still direct a 3'-end processing event in which the primary transcript is cleaved and then a poly (A) tail is added to the 3'-end product (Humphrey *et al.*, 1994). In the *ura4* gene of *S. pombe*, mutations in the polyadenylation signal not only abolish *in vivo* processing of the transcript but also affect transcription termination (Birse *et al.*, 1997). However, in other *S. pombe* genes the link is not so clear. In case of *nmt1* and *nmt2* genes, transcription extends for several kilobases beyond the poly (A) sites, despite the presence of efficient polyadenylation signals (Hansen *et al.*, 1998). These results suggest that some other elements are required in addition to the polyadenylation signal, whose presence, spatial arrangement and inherent strength correlates with the efficiency of termination. Eventually, these elements named as downstream elements (DSEs) with its transcriptional pausing activity have been demonstrated to play an important role in the termination event. DSEs have been defined in two of *S. pombe* genes, *ura4* and *nmt2*, and found that although they lack sequence homology, both function in the same locations and are orientation specific. Both of these elements are composed of multiple, degenerate sequence elements that act together to induce full polymerase pausing (Aranda and Proudfoot, 1999). These results indicate that pausing elements could be a common feature in genes of fission yeast and are quite indispensable for the precision of RNA polymerase termination. If that holds as an universal feature in *S. pombe* transcription termination, then there would obviously be DSE equivalents in the terminator region dissected out from the intronic region of *AmA1* genomic...
Transcription Termination of AmA1 Genomic Clone in S. pombe

close. The study tried to delineate the sequence contexts of these elements that are
recognized by S. pombe machinery in the heterologous genetic environment.

To demonstrate how precisely the termination event took place, a study was carried out using
a highly sensitive RT-PCR approach, firstly, to establish the extent of transcription past the
genes poly(A) site in the intronic region picked up in S. pombe. Almost a similar procedure
has been previously employed to define the region of transcription termination for the ura4
gene in S. pombe (Humphrey et al., 1994). The results revealed that RNA transcripts were
detectable at significant levels following the poly(A) site becoming undetectable beyond
~400bp downstream. It could thus be inferred that transcription terminated mostly close to
the 3' end processing site. This observation is consistent with the notion that there may be
tight linkage between polyadenylation and termination in yeast, as is being suggested by
various other investigators (Birse et al., 1997, 1998, Masatomo and Proudfoot, 1999). Signal
intensity measurement equivalent to the RT-PCR amplification, which in turn was
proportional to the amount of transcripts present, revealed that the 3' ends of these primary
transcripts were heterogeneous and amounts varied appreciably. However, detection of the
precise site(s) of transcription termination using RT-PCR mediated technique may not be
possible as it is basically a steady state transcript analysis. It is probable that RT-PCR
analysis may either be not sensitive enough to detect the unstable primary transcripts, or that
the assay may actually be measuring minor species of processed mRNA transcripts, present
at low levels.

The process of RNA Pol-II transcription termination has been proposed to involve
recognition of polyadenylation signal and communication of this to the elongating
polymerase (Logan et al., 1987; Connelly and Manley, 1988; Proudfoot, 1989). While the
mechanism employed to convey this information to the polymerase complex is unknown,
slowing down of the elongation rate of the polymerase after the poly (A) signal recognition
effectively retaining the transcribing complex close to the processing signal, is probably
instrumental in this interaction. Furthermore, lowering the levels of nascent transcription
beyond the poly(A) site would reduce the possibility of the elongating polymerase interfering
with the level of expression of genes located downstream (reviewed by Eggermont and
Transcription Termination of *AmA1* Genomic Clone in *S. pombe*

Proudfoot, 1993). Indeed, the possibility that transcriptional pausing may be associated with Pol-II termination in mammalian genes has been demonstrated (Enriquez-Harris *et al.*, 1991).

It is quite well known that transcriptional termination occurs downstream of 3' end processing events and the primary termination products are highly unstable. Consequently, they cannot be mapped by steady state RNA analysis techniques. To circumvent the limitations of steady state transcript analysis, transcription run-on (TRO) assay was employed to map polymerase density at the downstream region of poly(A) site. TRO analysis, in which nascent RNA is radiolabeled in isolated nuclei or permeabilized whole cells, gives a snapshot of the position of polymerases on any part of a gene and hence, a preferred method for termination studies. The conditions used in the assay permit transcriptionally engaged polymerases to elongate a short distance and incorporate radioactive label (Birse *et al.*, 1997, Hansen *et al.*, 1998). Signals specific to a particular probe in the TRO blot can be directly correlated with polymerase density over the region specific to the probe and hence, it can be used as a kinetic measure of polymerase processivity. Surprisingly, TRO profile revealed polymerase density over a large span of the genomic clone downstream to the poly(A) site. Polymerases transcribed for distances in excess of ~1.2 kb beyond the poly(A) site as was seen in this study, is analogous to the situation observed in case of *nmt1* and *nmt2* genes.

The high polymerase density observed over the intronic regions of *AmA1* genomic clone in the TRO assay is indicative of transcriptional pausing over those sequences that some of them might act as DSEs. Delineation of the probes which yielded high intensity signals (top three), to the DNA sequences they cover, points out two distinct regions, one spanning from 338 to 590 and the other from 589 to 949, while the third indicated the region 424-637, that overlaps with both. The signals represent the elevated level of polymerase density over these regions where elongating polymerases probably accumulate in response to transcriptional pause signals in a way that is characteristic of DSE. Furthermore, previously, northern results demonstrated the ability of these regions to promote 3'-end formation at upstream signals in consistence with the DSEs operating as a transcriptional pause signal. The DSE does not mediate polyadenylation or transcription termination in isolation, but rather appears
to act in concert with an upstream 3' end processing signal to mediate efficient termination of pol II transcription. The fact that transcription termination is produced far away from the poly(A) site and the putative DSEs is intriguing. Probably an interplay between multiple factors and sequence elements involves in bringing the RNA Pol-II transcription to a halt. The distance between elements and their sequence contexts may be important for the efficacy of the event. A working model for transcription termination in *S. pombe* has been proposed to fit the observations in the present investigation and to show the functionality of the intronic region of *AmAl* genomic DNA as transcription terminator, drawing parallels with the modular architectures available in other systems (Figure 15).

With an understanding of various *cis* elements involved in the termination event, it was of interest to see whether there is a unique structural signature present in the transcript which might have a role in the mechanistic of the process. RNA fold analysis of the terminator sequences revealed three distinct stem-loop structures arranged in tandem in the region spanning from position 245 to 367. The region was important keeping in view of the fact that it houses the SDE(s) and poly(A) site(s) and has primarily been implicated in the termination event. Though from the present set of experiments no role as such could be assigned to these stem-loop structures, which probably falls upstream to the DSEs, they can not be denied any functionality as well. It gains importance in the context of bacterial termination. At the very fundamental level, transcription as a process is very similar in bacteria and eukaryotes. The chemistry of RNA polymerization is identical in all types of organism, and the three eukaryotic RNA polymerases are each structurally related to the *E. coli* RNA polymerases. This means that the contact between an eukaryotic polymerase and its template DNA are probably very similar to the interaction described for bacterial transcription and the basic principle that the transcription is a step-by-step completion between elongation and termination also holds. These kinds of structural signatures are common to both intrinsic and factor dependent terminators occurred in bacteria as described previously (Brown, 1999). The possible involvement of a 'rho-like' helicase function in the pol II termination in yeast has been raised by the ability of the *E. coli* rho protein to mediate efficient RNA pol II transcription arrest *in vitro* (Wu and Platt, 1993). The occurrence of the stem-loop structures in tandem in a region which plays a crucial role in the 3' end processing in eukaryotes might
Animals

5'\[\text{AAUAAA} \rightarrow \text{DSE}\] 10 - 30 nts

S. cerevisiae

5'\[\text{EE} \rightarrow \text{PE}\] 16 - 27 nts

Plants

5'\[\text{FUE} \rightarrow \text{NUE(s)}\] 10 - 30 nts

S. pombe

5'\[\text{SDE1} \rightarrow \text{SDE2 / EE} \rightarrow \text{DSE}\]

AmAI genomic DNA

Northern Analysis

5'\[\text{SDE1} \rightarrow \text{SDE2} \rightarrow \text{DSE} \rightarrow \text{DSE2} \rightarrow \text{DSE3}\]

TRO Analysis

5'\[\text{Read-through transcript} \rightarrow \text{Prematurely terminated transcript (~350 nt)} \rightarrow \text{Prematurely terminated transcript (~500 nt)}\]

Figure 15. A. Modular architecture of cis elements involved in mRNA 3'-end processing. Schematic representation of the components of involved in 3'-end processing in animals, S. cerevisiae, plants and S. pombe. DSE, downstream element; EE, efficiency element; PE, positioning element; NUE, near upstream element; FUE, far upstream element. Representative sequence motifs characteristic of each class of elements are indicated wherever known. The cleavage/polyadenylation site is indicated by arrowhead. The drawing is not to scale. B. Proposed model of AmAI intronic sequence as transcription terminator in S. pombe. In the lower panel, salient points of two of the experiments which contributed in framing the model.
have a role in providing the structural requisites to stall an elongating polymerase at DSEs and in facilitating a molecular cross-talk between the components of 3' end processing and the polymerase.

The mechanism of pre-mRNA 3'-end formation and many of the factors involved in catalyzing these reactions are very similar between higher eukaryotes and S. cerevisiae (Manley and Takagaki, 1996; Keller and Minvielle-Sebastia, 1997). Because the efficiency of mammalian pre-mRNA 3'-end formation has been correlated with the degree of transcription termination, 3'-end formation factors have been implicated in the coupling of pol II termination with RNA processing. Indeed, the cleavage factors link transcription termination of RNA polymerase II with pre-mRNA 3'-end processing (Birse et al., 1998). With the information available on the pRAG transcription, the role of 3' end processing in transcription termination can not be conclusively demonstrated. Probably a TRO analysis with the polyadenylation signal mutants will clearly demonstrate linkage between this site with transcription termination. It was quite clear that on the putative SDE mutation, amount of read-through transcripts increased in case of pRABB and pRABN and what was more interesting was that in case of the latter, the amount of the aberrant spliced product was also increased appreciably. Again, the semi-quantitative RT-PCR analysis on read-through transcripts detected transcripts extended up to position 680, though the northern hybridization demonstrated the size of the transcripts not beyond ~600 nt. In light of all these observations it can be concluded that the transcripts must not have been cleaved at the poly(A) site(s) immediately as was detected in the downstream region, suggesting that there is a kinetic lag between transcription over the poly(A) signal and its effect on transcription termination. A kinetic lag seems to be essential for a fruitful interaction between factors involved in the cleavage at the poly(A) site and the polymerase. Recent evidences suggests that the CTD of RNA pol II acts as an important cofactor in the cleavage process (Hirose and Manley, 1998).

Poly(A) site mapping of AmAJ transcript in plant revealed a sequence 5'-CA-3' as the polyadenylation site at position 2607 of the gene with ATG at position +1. The sequence falls within the configuration of eukaryotic polyadenylation site consensus Y(A)n, generally located 10-30 nt downstream of the polyadenylation signal. While looking for sequences
Transcription Termination of AmA1 Genomic Clone in S. pombe

which match with the eukaryotic polyadenylation signal consensus, a sequence AATAAAA located 98 nt upstream of the poly(A) site has been detected. Although this very sequence can be viewed as the putative signal for polyadenylation in plant, further work in that direction is required to convincingly prove the proposition. On the other hand the RT-PCR studies with the AmA1 cDNA clone in S. pombe brought speculation that the polyadenylation in its transcripts took place in the R1044 binding site, i.e., at a sequence around position 2540 of the clone. This finding was in conformity with the northern result which indicated a transcript size of ~1.1 kb for the clone. Keeping in view that the transcript included 69nt from the nmtl promoter, polyadenylation might not have taken place in the nmtl terminator, which additionally would have added few hundreds of nucleotides to the transcript adding up the size to a larger one. Had it been the case, i.e., in the event of polyadenylation taking place in the nmtl terminator, the two different primers used in the semiquantitative RT-PCR would have yielded amplification of the same order. If that is the case, the findings with both the clones could be summed up to state that in S. pombe, the sequence element AATAAAA worked effectively as polyadenylation signal and polyadenylation took place ~25nt downstream, a distance conforming well to the eukaryotic average. In plant as well, if the same sequence is picked up as polyadenylation signal, polyadenylation at a comparatively far off site (~98nt) is intriguing. Probably the mechanism of 3' end maturation is conserved throughout the eukaryotic kingdom. The deviation from the basic realm of the event could be explained with an attribute of flexibility to the basic mechanism, which is justified keeping in view the diversity they possess in life forms. It is quite evident that in the terminator sequence, all the cis elements (SDE, poly(A) site, DSE etc.) required 3' end processing and transcription termination in S. pombe are present. The occurrence of a repertoire of diverse sequence elements in a heterologous piece of DNA could not be satisfactorily explained as a chance coincidence. Rather, an appreciation of the process from an evolutionary perspective sounds more plausible keeping in view of phylogenetic association of both the systems, S. pombe and plants. In that case, which prevents a termination event to take place in the intronic region in plant in line with the premature termination event in S. pombe, needs a second look over the proposition just stated.
Previous works with mammalian system show that a pause signal which was involved in RNA polymerase II transcription termination, not only enhances the utilization of an upstream poly(A) site which is otherwise out-competed by a stronger downstream poly(A) site but also rescues a poly(A) site that is inactive due to its location within an intron (Enriquez-Harris et al., 1991). But, splicing out-competes polyadenylation even at a strong poly(A) site (Levitt et al., 1989). Evidences accumulated over the years suggest that the mechanisms of 3' end processing and splicing are fundamentally similar across all eukaryotes, only specific adaptations which have evolved in different organisms bring differences (Rothnie, 1996). In *S. pombe* AmA1 genomic clone did not bring forth splicing, probably because the splicing machinery could not recognize properly the cis elements present in it. Polyadenylations at poly(A) sites in the intron got a chance to out-compete splicing. Precisely, the opposite might have occurred in case of the native system and splicing scored its due precedence out-competing an equally likely polyadenylation event in the intronic region downstream. Probably, a study with mutants which prevents splicing to occur in AmA1 gene would bring an answer to this conjecture.

3.5. Conclusion

*AmA1* genomic DNA when introduced in *S. pombe* underwent a premature termination event giving major and minor polyadenylated products. Poly(A) site and polyadenylation signal of the major product had been mapped and was found to match with their respective eukaryotic consensus sequences. A bipartite signal which involved in 3' end processing and characteristic of *S. pombe* genes was delineated in the AmA1 intronic sequence. Vertebrate poly(A) signal consensus AAUAAA found to work in *S. pombe* environment with certain variability. A sequence of size 376 bp, sufficient for a successful termination event had been dissected and cloned and its functionality as transcription terminator in *S. pombe* demonstrated. Three different regions had been delineated as candidate DSEs involved in polymerase pausing and bringing up of efficient 3' end processing at upstream poly(A) sites.
RNA structural analysis indicated three unique stem-loop structures, arranged in tandem in the region which was involved in 3' end processing. Poly(A) site of the AmA1 transcript in plant was mapped and the sequence context analyzed. The mechanisms of 3' end processing and transcription termination might be fundamentally the same across all eukaryotes and the little deviations could be interpreted as specific adaptations which have evolved in different organisms.