The EMBO Journal vol.7 no.5 pp.1497 – 1502, 1988

The mouse ribosomal gene terminator consists of three functionally separable sequence elements

Anne Kuhn, Andrea Normann, Ingrid Bartsch and Ingrid Grummt

Institut für Biochemie der Universität Würzburg, Röntgenring 11, 8700 Würzburg, FRG

Communicated by I.Grummt

The structural requirements for 3′ end formation of mouse pre-rRNA have been studied. Three sequence elements are shown to be required for accurate and efficient transcription termination by RNA polymerase I (pol I) assayed both in a cell-free transcription system and *in vivo* after transfection of rDNA minigene constructs into 3T6 cells. The essential termination signal is the previously identified 18-bp conserved element (AGG-TCGACCAG$^{AT}_{TA}$NTCCG) that contains a *Sal*I restriction site. This sequence motif (the 'Sal box') interacts with a specific nuclear protein that directs transcription termination. Here we demonstrate that the 'Sal box' sequence motif is sufficient for termination of pol I transcripts and the release of the nascent RNA chains from the template. However, in addition to this termination signal, pyrimidine-rich sequences flanking the box at the 5′ and 3′ side play a role in the efficient and correct formation of authentic pre-rRNA termini. Downstream sequences contribute to the efficiency of the termination reaction, whereas the position of 3′ end formation (i.e. 21 bp upstream of the 'Sal box') is affected by 5′ flanking regions. These flanking regions are recognized by at least two different nuclear factors which specifically bind to DNA sequences located upstream and downstream of the 'Sal box'.

Key words: rDNA/cell-free transcription/termination/pol I/ sequence elements

Introduction

Transcription initiation of genes coding for rRNA is a highly regulated and well-defined process. It involves the interaction of multiple proteins with defined sequence elements upstream of the transcription initiation site (for review see Jacob, 1986; Sollner-Webb and Tower, 1986). In addition to DNA–protein interactions, specific protein–protein interactions are required for the assembly of stable transcription complexes, the formation of which is a prerequisite for faithful transcription initiation (Wandelt and Grummt, 1983; Miesfeld and Arnheim, 1984; Iida *et al.*, 1985). Although important promoter domains and essential nucleotides within the core promoter have been identified (Grummt, 1982; Learned *et al.*, 1983; Sollner-Webb *et al.*, 1983; Skinner *et al.*, 1984; Kishimoto *et al.*, 1985; Haltiner *et al.*, 1986) and some of the *trans*-acting factors have been characterized and purified (Learned *et al.*, 1985; Bateman and Paule, 1986; Clos *et al.*, 1986a,b; Tower and Sollner-Webb, 1987), both the precise mechanism of transcription initiation by polymerase I (pol I) and the principles of regulation of pre-rRNA synthesis in response to the environmental conditions are not understood.

Even less is known about the events which cause termination of pre-rRNA transcription. In *Xenopus laevis* and in *Drosophila melanogaster*, the entire ribosomal gene repeat unit has been reported to be transcribed (Labhard and Reeder, 1986; McStay and Reeder, 1986; Tautz and Dover, 1986). Thus the 3′ end of the pre-rRNA in these species appears to be generated by a type of RNA processing rather than by a genuine termination event. In contrast, we have shown that in mammals RNA pol I terminates transcription at a defined point downstream of the 28S rRNA coding region (Grummt *et al.*, 1986a; Bartsch *et al.*, 1987). We have used gene fusions containing the mouse rDNA promoter and 3′ terminal spacer regions to investigate the mechanism of transcription termination by pol I. We could show that 3′ end formation of mouse pre-rRNA occurs 565 bp downstream of the 28S rRNA terminus 21 bp upstream of an 18-bp conserved sequence element AGG-TCGACCAG$^{TA}_{AT}$NTCCG, which is repeated eight times in the 3′ spacer between positions +587 and +1178 (Grummt *et al.*, 1985). This sequence element (the 'Sal box') mediates binding of a nuclear factor and transcription termination (Grummt *et al.*, 1986a). We have now examined whether this 18-bp sequence element by itself is sufficient or whether additional sequences contribute to the formation of correct 3′ ends. We find that the box itself does not constitute a fully efficient terminator but that the specific and efficient 3′ end formation of transcripts requires the presence of natural sequences upstream and downstream of the 'Sal box'.

Results

The sequence elements shown to be essential in vitro are also functionally active in vivo

In previous studies we have used a cell-free system and artificial ribosomal minigene constructs to demonstrate the importance of an 18-bp conserved sequence element containing a *Sal*I restriction site (the 'Sal box') in the process of transcription termination by mouse pol I and the involvement of a nuclear factor which specifically interacts with this termination signal. To investigate whether these DNA sequences play the same functional role within the cell, minigene constructs which contain parts of the bacterial marker gene coding for chloramphenicol acetyltransferase (CAT) positioned between the promoter and terminator region (Figure 1A), were transfected into 3T6 cells and the transcription of these gene fusions was monitored by analysis of cellular RNA. The presence of the bacterial gene fragment coding for CAT allowed us to detect pol I-specific transcripts derived from the exogenous DNA above the high background of cellular pre-rRNA. pMrC-BH is a plasmid

©IRL Press Limited, Oxford, England

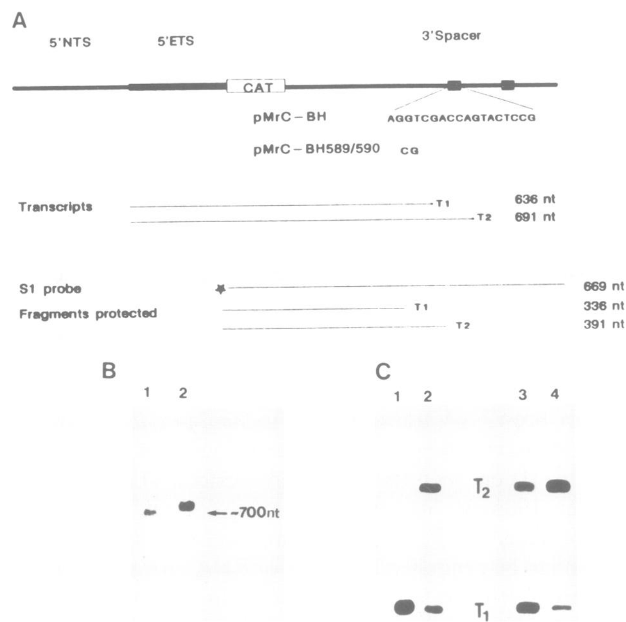

Fig. 1. Termination of rDNA minigene transcripts in the cell-free system and after transfection into 3T6 cells. **(A)** Schematic representation of the rDNA minigene constructs pMrC-BH and pMrC-BH589/590. Both plasmids contain 324 bp of the rDNA promoter region, 250 bp of the CAT-coding region and 378 bp of the 3′ terminal rDNA spacer. The boxes indicate the positions of the first and second 'Sal box', the sequence of which is shown below the first box. The two nucleotide exchanges introduced into positions +589 and +590 in the mutant clone are indicated. **(B)** Northern blot analysis of RNA from cells transfected with pMrC-BH or pMrC-BH589/590. Ten micrograms of RNA isolated from 3T6 cells transfected with pMrC-BH (**lane 1**) or pMrC-BH589/590 (**lane 2**) were resolved on an 1.5% agarose gel containing 8% formaldehyde. After blotting to nitrocellulose filters RNA was hybridized for 16 h at 42°C to the 1632-bp *Hin*dIII–*Bam*HI fragment derived from pSV2-CAT (Gorman, 1985). **(C)** S1 nuclease mapping of transcripts synthesized *in vitro* and *in vivo*. Transcripts synthesized in a 50-μl assay containing 30 ng of pMrC-BH and pMrC-BH589/590 respectively (**lanes 1** and **2**), or transiently expressed after transfection of 3T6 cells with these constructs (**lanes 3** and **4**) were hybridized to the 3′ labelled 669-bp *Pvu*II fragment from pMrC-BH and pMrC-BH589/590 respectively for 3 h at 64°C. After treatment with S1 nuclease the protected fragments were resolved on 6% sequencing gels. Transcripts terminated upstream of the first and second box (T_1 and T_2) yield a 336- and 391-nt protected fragment respectively.

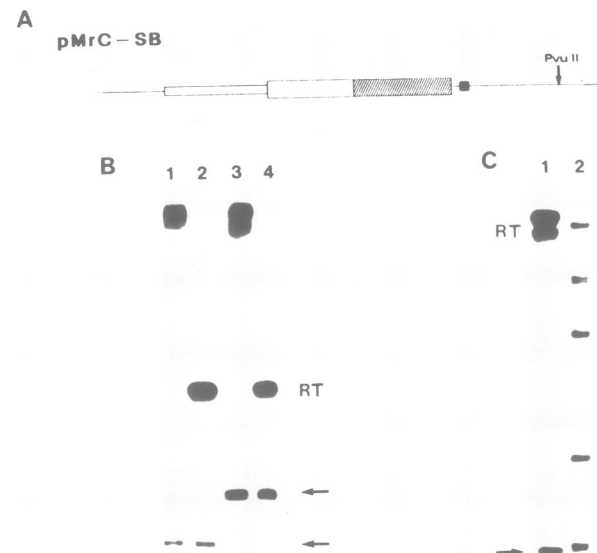

Fig. 2. Transcription termination by a 'Sal box' oligonucleotide. **(A)** Structure of the minigene plasmid pMrC-SB. The thin open bar represents 169 bp of the 5′ non-transcribed spacer (NTS), the thick open bar represents 135 bp of the 5′ terminal external transcribed spacer (ETS), the hatched region marks the 167-bp *Stu*I–*Pvu*II fragment from pSV2-CAT (Gorman, 1985), and the dark box indicates the position of the 18-bp synthetic oligonucleotide encompassing the 'Sal box' sequence. **(B)** Transcripts synthesized in the murine extract system from pMrC-SB (**lanes 1** and **2**) and pMrC552-650 (**lanes 3** and **4**). Thirty nanograms circular (**lanes 1** and **3**) and linear (**lanes 2** and **4**) template DNA were used in a 50-μl assay. The arrows mark the position of terminated transcripts. **(C)** 3′ end mapping of transcripts synthesized from pMrC-SB. Transcripts were synthesized in the cell-free system from pMrC-SB and were hybridized to 0.003 pmol (50 000 c.p.m.) of a labelled complementary RNA probe obtained by transcription of SP64C-SB by SP6 RNA polymerase (**lane 1**). Hybridization and RNase treatment was performed as described in Materials and methods. The RNase-resistant hybrids were run on a 6% sequencing gel along with size markers (pBR322/*Hpa*II, **lane 2**). The upper band corresponds to undigested probe and read-through transcripts (RT); the terminated transcripts are marked by an arrow.

covering the rDNA promoter region, 250 bp of the CAT-coding region and 378 bp of 3′ terminal spacer sequences from +335 to +712 including two termination sites. pMrC-BH589/590 is an identical construct which contains two base exchanges within the first 'Sal box' at positions +589 and +590. These point mutations have previously been shown to impair the interaction with the 'Sal box' binding protein which in turn results in inhibition of termination (Grummt *et al.*, 1986a).

Figure 1B shows CAT-specific transcripts transiently expressed in 3T6 cells transfected with pMrC-BH and pMrC-BH589/590 respectively. The Northern analysis reveals a distinct class of RNA molecules with an approximate length of 700 nt, a size which corresponds to the distance from the transcription start site to the termination site(s). The slightly slower mobility of the transcripts derived from cells transfected with the mutant pMrC-BH589/590 suggests that due to mutation of the first 'Sal box' the termination site T_2 located 55 bp further downstream has been used.

In order to map the 3′ ends of these cellular transcripts at the nucleotide level, nuclease S1 protection experiments were carried out (Figure 1C). Transcripts synthesized *in vitro* are shown for comparison. The 3′ ends of transcripts transiently expressed in 3T6 cells (lanes 3 and 4) are indistinguishable from RNA synthesized in the cell-free system (lanes 1 and 2). When pMrC-BH was used as template, the majority of transcripts mapped upstream of the first box (T_1) at position +565 with respect to the end of 28S rRNA (lanes 1 and 3). This site corresponds to the 3′ end of cellular pre-rRNA as determined before (Grummt *et al.*, 1985). In addition, transcripts terminated at the second box (T_2) are observed in both systems. The point mutations within the first 'Sal box' element contained in pMrC-BH589/590 strongly decreased termination at site T_1 and resulted in a preferential usage of the intact site T_2 (lanes 2 and 4). Thus, the 'Sal box' signal sequence appears to serve the same function in 3′ end formation of pre-rRNA both *in vitro* and *in vivo*.

The 'Sal box' is necessary and sufficient to mediate transcription termination

In view of the functional importance of the 'Sal box' in transcription termination we have examined whether these

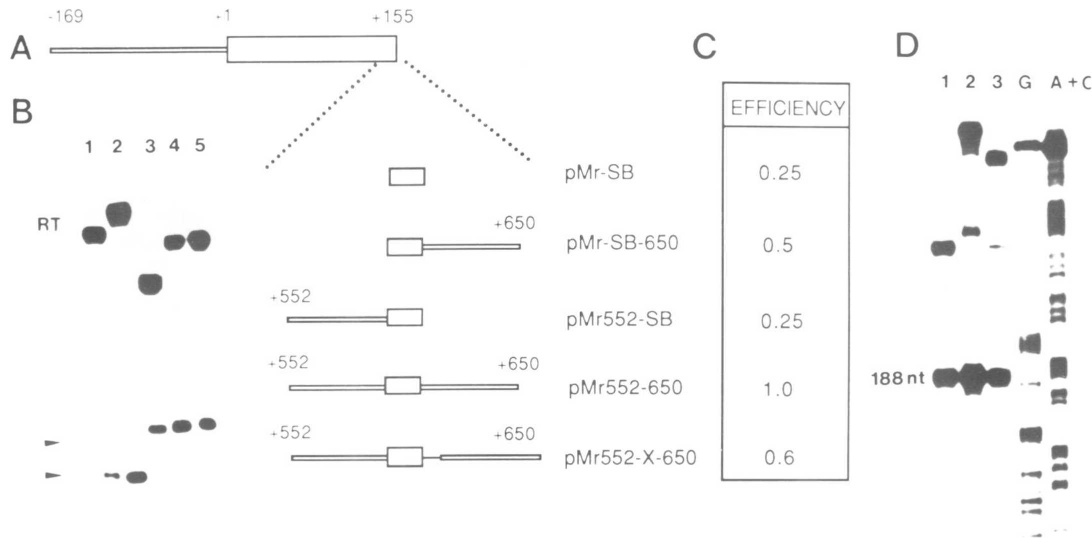

Fig. 3. Effect of adjacent DNA sequences on the efficiency and accuracy of transcription termination. (**A**) Schematic representation of the minigenes containing different 3′ terminal sequences. The promoter region (from position −169 to +155) to which the synthetic 'Sal box' (SB) oligonucleotide or 3′ terminal rDNA spacer fragments were fused is shown. The box marks the 18-bp 'Sal box' sequence. The flanking regions extend to position +552 and +650 respectively with respect to the end of the 28S rRNA coding region. (**B**) Transcripts derived from the different minigenes. The templates were cleaved with *Pvu*II and 30 ng of template DNA plus 70 ng pUC9 were used in the transcription assay. **Lane 1**, pMr-SB; **lane 2**, pMrSB-650; **lane 3**, pMr552-SB; **lane 4**, pMr552-650; **lane 5**, pMr552-X-650. The transcripts were analysed by electrophoresis on a 4.5% polyacrylamide gel. The lower RNA bands (marked by an arrowhead) represent terminated transcripts, the upper bands are read-through transcripts (RT). (**C**) Evaluation of the termination efficiency of different minigene constructs calculated from the relative amounts of terminated versus read-through transcripts. (**D**) 3′ end mapping of transcripts derived from minigenes containing different 3′ terminal flanking sequences. Thirty nanograms of circular template DNA pMrC552-SB (**lane 1**), pMrC552-650 (**lane 2**) and pMrC552-X-650 (**lane 3**) were transcribed *in vitro* and RNA was analysed by S1 mapping using a 287-bp *Hin*dIII−*Eco*RI fragment from pMrC552-650. The S1-resistant DNA fragments were resolved on a 6% polyacrylamide−8 M urea gel adjacent to chemical cleavage reactions of probe DNA. The 188-nt band corresponds to transcripts the 3′ ends of which map at position +565.

18 nucleotide pairs are sufficient to stop transcription, or whether they work in concert with additional elements of the rDNA spacer. A synthetic 18-bp oligonucleotide containing the 'Sal box' sequence was fused downstream of the rDNA promoter and a fragment from the CAT gene to yield the recombinant plasmid pMrC-SB (Figure 2A). The transcripts derived from this artificial minigene construct in the cell-free transcription system are shown in Figure 2B. In the circular form (lane 1) this template yields a defined class of RNA molecules whose length corresponds to the distance from the initiation site to the synthetic 'Sal box' oligonucleotide. Read-through transcripts accumulate on top of the gel. After truncation with *Pvu*II, run-off transcripts 532 nt long (RT) are synthesized in addition to terminated RNA chains (lane 2). The efficiency of transcription termination at the synthetic sequence motif, however, was significantly lower compared with the control pMrC552-650, an identical construct containing natural adjacent DNA sequences (lanes 3 and 4).

To locate precisely the 3′ ends of RNA terminated at the synthetic 'Sal box', trancripts synthesized from pMrC-SB were mapped by the RNase protection technique. From the Maxam−Gilbert sequencing reactions run in parallel to the probes (not shown), the 3′ ends of the terminated transcripts can be localized at the nucleotide level. Surprisingly, the 3′ terminus of transcripts mapped only 11 bp upstream of the box, whereas transcripts from constructs which contain natural termination sites have been shown to map 21 bp upstream of the 'Sal box' at nucleotide +565. This result suggests that the pyrimidine-rich sequences which normally flank the box may affect the formation of correct 3′ termini.

Sequences flanking the 'Sal box' at the 5′ and 3′ site determine the accuracy and efficiency of 3′ end formation

To study which sequences determine the specificity of 3′ end formation and affect the efficiency of the termination process, rDNA minigenes were constructed which contained in addition to the 'Sal box' signal (from +587 to +604) adjacent 3′ spacer regions encompassing upstream sequences from position +552 and downstream sequences extending to position +650 respectively. The recombinant plasmids used and the transcripts generated in the cell-free system at linear templates are shown in Figure 3A and B. The lower band represents terminated RNA molecules, the upper band shows read-through transcripts. The ratio of terminated versus read-through transcripts is a measure for the efficiency of the termination reaction. In this experiment the assay conditions were chosen (template concentrations and extracts used) to yield an approximate 1:1 ratio of terminated to read-through transcripts in the control construct containing 3′ terminal sequences from +552 to +650 (pMr552-650). The quantitative evaluation of the termination efficiency (Figure 3C) demonstrates that the presence of sequences flanking the 'Sal box' at the 3′ site augments termination ~2-fold (pMrSB-650). Addition of 5′ flanking regions has no effect on the efficiency of termination (pMr552-SB). If both 5′ and 3′ adjacent sequences were contained in the minigene construct (pMr552-650) the termination efficiency was increased above the level observed with 3′ flanking sequences alone (pMrSB-650), indicating that both regions contribute to the efficiency of the termination reaction.

Fig. 4. Release of terminated transcripts from the template. 60 ng of minigene constructs pMrC-SB (**lanes 1–3**) or pMrC552-650 (**lanes 4–6**) were transcribed in a 100-μl transcription assay in the presence of 140 ng of pUC9 DNA. After incubation for 60 min at 30°C, 50 μl of the reaction mixture was centrifuged at 15 000 g for 10 min. Radiolabelled RNA from the uncentrifuged reaction (**lanes 1** and **4**), the resultant supernatant (**lanes 2** and **5**) and the pellet fraction (**lanes 3** and **6**) was extracted and resolved on 4.5% polyacrylamide gels.

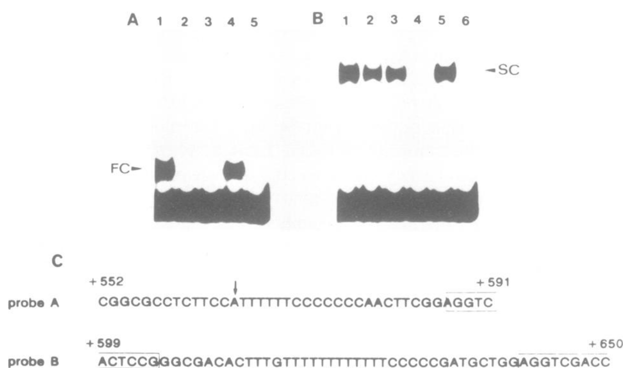

Fig. 5. Electrophoretic mobility shift assay with DNA fragments containing DNA sequences upstream or downstream of the first 'Sal box'. (**A**) Mobility shift with the 85-bp *Eco*RI–*Hin*dIII fragment from pUC552-591. Each assay contained ~20 000 c.p.m. (~0.5 ng) of 3′ end-labelled DNA, 15 μg of nuclear extract protein, and 500 ng of different competitor DNAs: pUC9 (**lane 1**), pUC552-591 (**lane 2**), pUC552-598 (**lane 3**), pUC599-650 (**lane 4**) and pUCT$_1$ (**lane 5**). The fast-migrating DNA–protein complex (FC) characteristic for binding of protein(s) to 5′ flanking regions is marked. (**B**) Mobility shift with the 71-bp *Eco*RI–*Hin*dIII fragment from pUC599-650. The assay is analogous to that of (A). The following competitor DNAs were used: pUC9 (**lane 1**), pUC552-591 (**lane 2**), pUC552-598 (**lane 3**), pUC599-650 (**lane 4**), pUCT$_1$ (**lane 5**) and pUCT$_2$ (**lane 6**). The more slowly migrating DNA–protein complex formed with the 3′ flank is designated SC. (**C**) Nucleotide sequence of the regions flanking the 'Sal box' motif. The sequence of the non-coding strand is shown. The arrow marks the position of the 3′ end of 45S pre-rRNA at position +565. The boxed nucleotides represent part of the 'Sal box' element.

Next we examined whether the effect of the 3′ flanks relies on a spatial relationship between the elements themselves. Clone pMr552-X-650 is identical to clone pMr552-650 except that 13 bp foreign DNA from the pUC9 multiple cloning site were inserted between positions +604 and +605, thus expanding the distance between the 'Sal box' and the 3′ flank. Since the efficiency of transcription termination is significantly decreased by this insertion, we suggest that both sequence elements—the 'Sal box' and neighbouring sequences—function synergistically and that their effect is position dependent.

In order to determine whether the 3′ ends of transcripts derived from templates with 5′ flanks map 21 bp upstream of the box as found for cellular pre-rRNA, S1 mapping experiments were performed. As shown in Figure 3D, the termini of transcripts synthesized from all three templates containing 5′ flanking sequences map precisely at +565, the position which corresponds to the 3′ end of cellular pre-rRNA. In the absence of these upstream sequences RNA chains extending 10 nt further downstream are formed (see Figure 2). Thus the generation of correct 3′ ends upstream of the box is affected by the presence of 5′ flanking sequences.

Transcripts terminated at the 'Sal box' are released from the template

In order to get insight into the function of the neighbouring sequences we studied whether the flanks facilitate the release of the nascent RNA from the DNA. Two minigene constructs pMrC-SB and pMrC552-650, which differ by the presence of sequences flanking the 'Sal box', were used as templates in the cell-free system. After transcription the reaction assay was fractionated into supernatant and pellet by centrifugation. By this procedure the DNA together with associated proteins which have assembled into large sedimentable complexes is quantitatively pelleted (Culotta *et al.*, 1985). By analysis of the nucleic acids present in the resulting pellet and supernatant fraction it is possible to estimate the relative amount of RNA that is bound to the template versus the number of RNA molecules liberated from the DNA. As shown in Figure 4, the large majority of minigene transcripts terminated at the box are recovered in the soluble fraction (lanes 2 and 5), irrespective of whether or not flanking sequences are contained in the template. Long non-terminated transcripts which have traversed the closed circular template are predominantly found in the sedimentable fraction, indicating that they were bound to the DNA. This finding demonstrates that at the 'Sal box' the nascent RNA chains are released from the template both in the presence and absence of flanking sequences. Thus the box itself acts as a terminator which makes the pol I stop and release the RNA from the elongation complex.

Distinct proteins bind to DNA sequences upstream and downstream of the 'Sal box' sequence motif

In view of the functional importance of the sequence flanking the 'Sal box' we investigated whether these neighbouring regions are recognized by specific DNA binding proteins. For this we used the gel retardation assay which was initially described by Fried and Crothers (1981) and by Garner and Revzin (1981). Two fragments encompassing sequences from +552 to +591 and from +599 to +650 respectively were 3′ end-labelled and incubated with nuclear extract proteins in the presence of different competitor DNAs. The DNA–protein complexes were separated from free DNA by electrophoresis in low-ionic-strength polyacrylamide gels. As shown in Figure 5, different complexes were formed depending on the DNA fragment used. There was a fast migrating complex (FC) which is specific for sequences located upstream of the 'Sal box' (Figure 5A). This specific protein–DNA complex can be competed by homologous sequences contained in the plasmids pUC552-591 (lane 2), pUC552-598 (lane 3) and pUCT$_1$ (lane 5). Formation of this fast complex is not impaired by the presence of pUC9 (lane 1) or pUC599-650 (lane 4).

Apparently, the sequences downstream of the box are recognized by different proteins. The DNA–protein complex formed with the 3′ flanks (SC) migrates more slowly

than that formed with the 5′ flanking DNA probe (Figure 5B). The sequence specificity of the complex formed with the 3′ flank is clearly distinct from that of the complex formed with the upstream sequences. Complex formation is not prevented by the presence of pUC9 (lane 1) and the subclones which contain 5′ flanking sequences between positions +434 and +598, i.e. pUC552-591 (lane 2), pUC552-598 (lane 3) and pUCT$_1$ (lane 5). The formation of the specific complexes is competed by sequences located downstream of the box present in pUC599-650 (lane 4) and pUCT$_2$ (lane 6). Thus the pyrimidine-rich regions flanking the box at both sides interact with different proteins which exhibit different sequence specificities. This is an unexpected finding, taking into account the sequence similarity of the two fragments (Figure 5C). In view of the functional differences of 5′ and 3′ flanking sequences on the accuracy and efficiency of transcription termination, it is suggested that these effects are mediated by different proteins which bind specifically to these neighbouring sequences.

Discussion

We have investigated the sequence requirements for 3′ end formation of mouse pre-rRNA. In accord with our previous findings we demonstrate that the most important element is the 18-bp 'Sal box' sequence motif which functions as termination signal and which on its own is capable of stopping transcript elongation and releasing the nascent RNA chains. However, transcription termination brought about by the synthetic 'Sal box' removed from its normal gene environment was much less efficient compared with the natural terminator containing 5′ and 3′ adjacent sequences. More importantly, the 3′ ends of transcripts which in the presence of 5′ flanks map 21 bp upstream of the box, extend 10 nt further downstream in the absence of natural flanks. Thus, the neighbouring sequences contain information which augments the termination reaction and plays a role in the positioning of 3′ termini.

In previous reports we have presented several lines of experimental evidence suggesting that the 'Sal box' is a transcription terminator and not a processing signal (Grummt *et al.*, 1986b). However, at present we cannot exclude the possibility that 3′ end formation 21 bp upstream of the box is the result of a limited 3′ terminal trimming or processing reaction which is coupled to transcription termination. In this case pol I would stop transcription in front of the box at nucleotide +575 (the stop observed with the template pMrC-SB containing the synthetic 18-bp sequence), followed by a very rapid cleavage of the last 10 nt. In fact, such putative processing intermediates have been observed both *in vivo* and in the cell-free transcription system by pulse-chase experiments (A.Kuhn, unpublished results).

It is very likely that both the stimulating effect of the 3′ flank and the positioning effect of the 5′ flank are mediated by *trans*-acting factors which recognize these sequences. Using the mobility shift assay we have shown that the flanking regions interact specifically with different proteins. Both the nature of these proteins and their function in the process of 3′ end formation is still unknown. We also do not know whether there is a functional interaction of the flank binding proteins with the nuclear factor which binds to the 'Sal box'. The fact, however, that a 13-bp insertion between the box and the 3′ flank largely suppresses the stimulatory effect of the downstream region suggests that the spatial array of both sequence elements is a prerequisite for their synergistic action.

Such a synergistic effect of different sequence elements in transcription termination is not without precedent. The *X.laevis* ribosomal gene terminator (the T3 box) represents a 12-bp element which is required to specify the formation of correct 3′ ends (Labhard and Reeder, 1987). Removal of the T3 box from its normal gene environment results in a significant amount of read-through transcripts, indicating that additional sequences may be required for complete terminator function. Similarly, it has been shown that a conserved box in the 3′ flanking region of U1 and U2 genes is absolutely required for precursor 3′ end formation. This box acts in concert with two upstream sequence elements which affect the efficiency and accuracy of mature 3′ end formation (Hernandez, 1985; Mattaj and De Robertis, 1985; Yuo *et al.*, 1985; Ciliberto *et al.*, 1986). Furthermore, two or three distinct elements have been shown to be required for efficient generation of 3′ ends of both polyadenylated and non-polyadenylated mRNA termini. A 35-base region, located downstream of the AAUAA sequence, the essential signal for 3′ terminal processing of the primary transcript (reviewed by Birnstiel *et al.*, 1985), has been shown to be essential for rabbit β-globin formation (Gil and Proudfoot, 1987). In the sea urchin H2A gene tripartite sequences including part of the coding region together with two separate sequences in the 3′ flanking region constitute an RNA pol II termination signal (Johnson *et al.*, 1986). However, there is a marked difference in the molecular mechanism of 3′ end formation of different classes of RNA. Whereas the generation of mRNA 3′ termini is brought about by a processing event which involves the specific interaction of defined proteins with sequences of the mRNA precursor, pre-rRNA transcription termination clearly involves DNA–protein interactions.

Materials and methods

Plasmid constructions

All minigene constructs used are derivatives of pMrWT (Skinner *et al.*, 1984). pMrWT contains a 324-bp *Sal*I–*Sma*I fragment (from −169 to +155) from the 5′ region of the rDNA repeat. Different 3′ terminal spacer sequences are contained in pMr-SB, pMr552-SB, pMrSB-650 and pMr552-650. The numbers mark the boundaries of flanking sequences present in the gene fusions, SB marks the 'Sal box' sequence motif extending from position +587 to +604 relative to the 3′ end of 28S rRNA. Thus, in pMr-SB a synthetic oligonucleotide containing the 18 bp of the first 'Sal box' is fused downstream of the rDNA promoter, pMr552-SB contains the box and upstream sequences starting at nucleotide +552. Similarly, pMrSB-650 contains the box and 3′ flanks extending to nucleotide +650. The terminator region in the normal gene environment covering 3′ spacing sequences from +552 to +650 is contained in plasmid pMr552-650. For construction of the spacing mutant pMr552-X-650, the plasmid pMr552-SB was cleaved with *Eco*RI and the 48-bp *Rsa*I–*Hin*dII fragment (from +599 to +646) was inserted by blunt-end ligation. Thus a recombinant DNA was formed (pMr552-X-650) which is identical to pMr552-650 except for a 13-bp insertion between the box and the downstream sequences. To use these kinds of constructs also in transfection experiments and to facilitate S1 mapping, a reporter gene fragment was inserted between the promoter and terminator region. For this, the minigenes were cleaved with *Ava*I and the 167-bp *Stu*I–*Pvu*II fragment derived from pSV2-CAT was inserted to yield the plasmids pMrC-SB, pMrC552-SB, etc.

The plasmid pMrC-BH and the point mutant pMrC-BH589/590 (see Figure 1A) are derivatives of pPTBH (Grummt *et al.*, 1985, 1986a). They contain the 324-bp 5′ terminal rDNA fragment, the 250-bp *Hin*dIII–*Pvu*II fragment from the CAT-coding region present in pSV2-CAT (Gorman, 1985), and the 378-bp *Bam*HI–*Hin*fI fragment from positions +335 to +712

relative to the 28S rRNA terminus. In pMrC-BH589/590 the guanine and thymine residues at positions +589 and +590 have been converted into C and G residues respectively (Grummt *et al.*, 1986a).

In the mobility shift assays subclones containing different regions from the 3′ spacer were used. pUC552-591 contains a *Fnu*DII−*Hin*dII fragment covering sequences from +552 to +591, pUC599-650 contains a *Rsa*I−*Sal*I fragment which extends from +599 to +650. pUC552-598 covers sequences from the *Fnu*DII site (at position +552) to the *Rsa*I site at +598. The plasmids $pUCT_1$ and $pUCT_2$ used as competitors contain the 171-bp (position +434 to +604) and 83-bp (from +604 to +686) *Nci*I fragments respectively, cloned into the *Sma*I site of pUC9 (Grummt *et al.*, 1986a).

In vitro transcription assay

S-100 and nuclear extracts were prepared from cultured Ehrlich ascites cells according to Weil *et al.* (1979) and Dignam *et al.* (1983). For *in vitro* transcription 30 ng of template DNA and 70 ng of pUC9 carrier DNA were incubated in a 50-μl assay in the presence of a mixture of nuclear and S-100 extracts. The reaction mixture contained 12 mM Hepes (pH 7.9), 85 mM KCl, 0.12 mM EDTA, 5 mM $MgCl_2$, 10 mM creatine phosphate, 0.6 mM each of ATP, CTP and UTP, 12.5 μM GTP and 1 μCi of [α-^{32}P]-GTP. The mixture was incubated for 60 min at 30°C and was processed for gel analysis as previously described (Grummt, 1982). To prepare unlabeled RNA the assay was scaled up to 100 μl and transcription was performed in the presence of 0.6 mM of each of the four ribonucleoside triphosphates. Samples to be analysed by S1 mapping were treated with DNase I before phenol extraction and ethanol precipitation.

Transfection of cells and preparation of cellular RNA

Subconfluent cultures of 3T6 cells were transfected with 20 μg of supercoiled plasmid DNA by the calcium phosphate technique (Graham and Van der Eb, 1973). The cells were harvested 44−48 h after transfection. RNA was isolated by the method of Chirgwin *et al.* (1979), dissolved in sterile water at 2 mg/ml and used for hybridization.

Nuclease S1 mapping of 3′ ends of transcripts

RNA from a preparative (100 μl) *in vitro* transcription assay or 50 μg of cellular RNA extracted from transfected cells was mixed with the labelled hybridization probe, precipitated with ethanol and dissolved in 25 μl hybridization buffer (80% formamide, 0.4 M NaCl, 0.04 M Pipes, pH 6.4 and 1 mM EDTA). After hybridization for 3 h at the appropriate temperature (see figure legends) the hybrids were treated with 100 U S1 nuclease (Berk and Sharp, 1977). The hybrids were purified by extraction with phenol−chloroform and analysed on 6% sequencing gels along with size markers (pBR322/*Hpa*II) or a Maxam−Gilbert sequencing ladder of probe DNA.

RNase protection assay

A 201-bp *Hin*dIII fragment derived from pMrC-SB was fused in reverse orientation to the SP6 promoter to yield the recombinant plasmid SP64C-SB. 300 ng of this plasmid were linearized with *Eco*RI and transcribed in a 10-μl reaction mixture containing 40 mM Tris−HCl (pH 7.5), 6 mM $MgCl_2$, 10 mM DTT, 2 mM spermidine, 0.1 mg/ml BSA, 0.5 mM each of ATP, CTP and UTP, 25 μCi [α-^{32}P]GTP, 15 U RNase inhibitor and 5 U SP6 RNA polymerase. 0.003 pmol of this labelled complementary RNA probe was hybridized to cold RNA synthesized in a 25-μl *in vitro* assay containing 15 ng of pMrC-SB as template. After hybridization at 37°C in 80% formamide, 0.4 M NaCl, 0.04 M Pipes (pH 6.4) and 1 mM EDTA for 3 h, 300 μl of 0.3 M Na acetate (pH 7.0), 5 mM EDTA, 3 μg/ml RNase A and 7.5 U/ml RNase T_1 were added and the single-stranded RNA was digested for 1 h at 37°C. The reaction was terminated by the addition of 7 μl 10% SDS and 40 μg proteinase K and subsequent incubation for 15 min at 42°C. The RNA−RNA hybrids were extracted with phenol−chloroform and analysed on a 6% sequencing gel.

Electrophoretic mobility shift assay

The *Eco*RI−*Hin*dIII fragments from pUC552-591 (85 bp) or from pUC599−650 (71 bp) were labelled at the 3′ ends. Approximately 20 000 c.p.m. (~0.5 ng) of the labelled DNA was incubated in 12.5 μl of binding buffer [12 mM Hepes, pH 8.0, 100 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM dithiothreitol, 8% glycerol, 10 μg tRNA, 1 μg poly(dI-dC)] in the presence of ~20 μg of nuclear extract protein. After 15 min at room temperature, samples were loaded onto low-ionic-strength (30 mM Tris−HCl, pH 8.0, 30 mM boric acid, 1 mM EDTA) 6% polyacrylamide gels and electrophoresed at 10 V/cm for 2 h at room temperature. The gel was then dried and analysed by autoradiography.

Acknowledgements

We are grateful to Mrs E.Zimmer for typing the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft and the Fond der Chemischen Industrie.

References

Bartsch,I., Schoneberg,C. and Grummt,I. (1987) *Mol. Cell. Biol.*, **7**, 2521−2529.

Bateman,E. and Paule,M.R. (1986) *Cell*, **47**, 445−450.

Berk,A.J. and Sharp,P.A. (1977) *Cell*, **12**, 721−732.

Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349−359.

Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294−5299.

Ciliberto,G., Dathan,N., Frank,R., Philipson,L. and Mattaj,I. (1986) *EMBO J.*, **5**, 2931−2937.

Clos,J., Buttgereit,D. and Grummt,I. (1986a) *Proc. Natl. Acad. Sci. USA*, **82**, 604−608.

Clos,J., Normann,A., Öhrlein,A. and Grummt,I. (1986b) *Nucleic Acids Res.*, **14**, 7581−7595.

Culotta,V.C., Wides,R.J. and Sollner-Webb,B. (1985) *Mol. Cell. Biol.*, **5**, 1582−1590.

Dignam,J.D., Lebowitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **2**, 1475−1489.

Fried,M. and Crothers,D.M .(1981) *Nucleic Acids Res.*, **9**, 6505−6525.

Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047−3060.

Gil,A. and Proudfoot,N.J. (1987) *Cell*, **49**, 399−406.

Gorman,C.M. (1985) In Glover,D.M. (ed.) *DNA Cloning. A Practical Approach*. IRL Press, Oxford, Vol. II, pp. 143−190.

Graham,F.C. and Van der Eb,A.J. (1973) *Virology*, **52**, 456−464.

Grummt,I. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6908−6911.

Grummt,I., Maier,U., Öhrlein,A., Hassouna,N. and Bachellerie,J.-P. (1985) *Cell*, **43**, 801−810.

Grummt,I., Rosenbauer,H., Niedermeyer,I., Maier,U. and Öhrlein,A. (1986a) *Cell*, **45**, 837−846.

Grummt,I., Kuhn,A., Bartsch,I. and Rosenbauer,H. (1986b) *Cell*, **47**, 901−911.

Haltiner,M.M., Smale,S.T. and Tjian,R. (1986) *Mol. Cell. Biol.*, **6**, 227−235.

Hernandez,N. (1985) *EMBO J.*, **4**, 1827−1837.

Iida,C.T., Kownin,P. and Paule,M.R. (1985) *Proc. Natl. Acad. Sci. USA*, **81**, 299−303.

Jacob,S.T. (1986) *Mol. Cell. Biochem.*, **70**, 11−20.

Johnson,M.R., Norman,C., Reeve,M.S., Scully,J. and Proudfoot,N. (1986) *Mol. Cell. Biol.*, **6**, 4008−4018.

Kishimoto,T., Nagamine,M., Sasaki,T., Takakusa,N., Miwa,T., Kominami,R. and Muramatsu,M. (1985) *Nucleic Acids Res.*, **13**, 3515−3532.

Labhard,P. and Reeder,R.H. (1986) *Cell*, **45**, 431−443.

Labhard,P. and Reeder,R.H. (1987) *Mol. Cell. Biol.*, **7**, 1900−1905.

Learned,R.M., Smale,S.T., Haltiner,M.M. and Tjian,R. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3558−3562.

Learned,R.M., Cordes,S. and Tjian,R. (1985) *Mol. Cell. Biol.*, **5**, 1358−1369.

McStay,B. and Reeder,R.H. (1986) *Cell*, **47**, 913−920.

Mattaj,I.W. and De Robertis,E.M .(1985) *EMBO J.*, **2**, 1883−1891.

Miesfeld,R. and Arnheim,N. (1984) *Mol. Cell. Biol.*, **4**, 221−227.

Skinner,J.A., Öhrlein,A. and Grummt,I. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2137−2141.

Sollner-Webb,B. and Tower,J. (1986) *Annu. Rev. Biochem.*, **55**, 801−830.

Sollner-Webb,B., Wilkinson,J.A.K., Roan,J. and Reeder,R.H. (1983) *Cell*, **35**, 199−206.

Tautz,D. and Dover,G.A. (1986) *EMBO J.*, **5**, 1267−1273.

Tower,J. and Sollner-Webb,B. (1987) *Cell*, **50**, 873−883.

Wandelt,C. and Grummt,I. (1983) *Nucleic Acids Res.*, **11**, 3795−3809.

Weil,P.A., Luse,D.S., Segall,J. and Roeder,R.G. (1979) *Cell*, **18**, 469−484.

Yuo,C.-Y., Ares,M.,Jr and Weiner,A.M. (1985) *Cell*, **42**, 193−202.

Received on January 27, 1988; revised on February 15, 1988