Natural point mutations within rat rDNA transcription terminator elements reveal the functional importance of single bases for factor binding and termination

Milko B. Kermekchiev* and Ingrid Grummt

Institut für Biochemie, Röntgenring 11, 8700 Würzburg, FRG

Received March 13, 1987; Revised and Accepted April 23, 1987

ABSTRACT
The rat rDNA transcription unit extends 560-565 bp into the spacer downstream of the 28S rRNA coding region. The site of 3' end formation is located in front of a conserved 18 bp sequence element which is repeated eight times in the 3' spacer between nucleotides +582 and +1767 relative to the 3' terminus of 28S rRNA. These sequence motifs are almost identical to the RNA polymerase I transcription termination signal (the Sal I box) that has previously been identified in the 3' terminal spacer of mouse rDNA. Interestingly, each of the single rat elements contains one or more base substitutions as compared to the murine Sal I box. Individual rat Sal I boxes were cloned and tested for their ability to interact with the murine termination factor and to direct transcription termination. It is shown that five of the eight boxes represent genuine transcription terminators, while three elements contain certain point mutations which are not recognized by the nuclear Sal I box-binding protein and therefore are functionally inactive.

INTRODUCTION
Recent studies on transcription of ribosomal genes in several organisms have demonstrated that 3' end formation of ribosomal precursor RNA (pre-rRNA) is brought about by either transcription termination of RNA polymerase I (pol I) or by processing of long unstable transcripts. The latter case has been reported to occur in Xenopus laevis and Drosophila melanogaster (1, 2). There almost all of the spacer is transcribed as part of the primary transcript and thus the well-defined pre-rRNA molecules are the first semistable processing intermediates. In mouse, on the other hand, we could demonstrate that pol I terminates transcription 565 bp downstream of the 3' end of 28S rRNA (3). An 18 bp conserved sequence element (designated Sal I box) which is repeated several times in the 3' spacer, functions as a transcription terminator. It is recognized by a nuclear factor which in turn directs the stop of pol I movement along the rDNA template and the release of pre-rRNA (4).
Recently the nucleotide sequence of the 3' terminal spacer region of rat rDNA has been determined (5). A comparison of the rat sequence with the analogous region from mouse rDNA reveals a significant divergence in the nucleotide composition beyond the 28S rRNA coding region. Most interestingly, several copies of a conserved 18 bp sequence element structurally analogous to the murine Sal I box motif are found in the rat spacer. Since both the primary structure, the localization and the redundancy are very similar to the mouse termination signal, it is most likely that the rat and the mouse sequence elements are functionally equivalent. However, there is one remarkable difference between the Sal I box elements in both species. Whereas in mouse the sequence of the eight boxes is identical, each of the rat elements shows single or several base exchanges as compared to the mouse consensus sequence (see Fig. 3). The availability of these natural point mutations enables us to study the functional importance of defined nucleotides within this signal sequence. We show that the ability of the murine transcription termination factor to interact with the different rat Sal I box elements correlates with their efficiency to mediate transcription termination in a cell-free system.

**MATERIALS AND METHODS**

**Plasmid Constructions**

The plasmids pRT12, pRT123, pRT1-6, pRT7 and pRT8 originate from pRr22 (5) bearing a 2700 bp Eco RI-Bam HI rat rDNA fragment which contains 595 bp from the 28S rRNA coding region and 2105 bp from the 3' terminal spacer. As shown in Fig. 1 plasmids pRT12 and pRT123 contain the Sau 3A- Sal I fragments encompassing nucleotides +352 to +679 and +352 to +840 (relative to the 3' terminus of 28S rRNA), respectively. pRT1-6 contains the 756 bp Sau 3A fragment (+352 to +1108); pRT7 contains the 415 bp Sau 3A - Pvu II fragment (+1170 to +1585) and pRT8 contains the 444 bp Pvu II-Sau 3A fragment (+1586 to +2029). All fragments were cloned into pUC9. Mouse-rat minigenes (designated pMPRT) were constructed by fusing rat rDNA 3' spacer fragments to a 324 bp Sal I - Sma I fragment from the 5' terminal region of mouse rDNA covering sequences from -169 to +155 with respect to the initiation site. This promoter fragment fused to 3' terminal mouse rDNA spacer sequences (from +433 to +604) yielded the recombinant plasmid pMrT 1 (4).

**In vitro Transcription Assays**

Nuclear extracts were prepared from Ehrlich ascites cells according to Dignam et al. (6). For in vitro transcription 10-100 ng of template DNA were incubated in a 50 µl assay in the presence of 30 µl of nuclear extract. The
The assay contained 12 mM HEPES (pH 7.9), 85 mM KCl, 0.12 mM EDTA, 5 mM MgCl₂, 10 mM creatine phosphate, 0.6 mM each of ATP, CTP and UTP, 12.5 μM GTP and 1-2 μCi of (α-³²P)GTP (400 Ci/mmol). The mixture was incubated for 60 min at 30°C and RNA was analyzed as described previously (7).

**Exonuclease III Protection Experiments**

Exonuclease III protection assays were performed essentially as described by Wu (8). For each binding reaction about 4,000 cpm (about 2 ng DNA) of labelled fragment was incubated for 20 min at 30°C in 25 μl of binding buffer (12 mM HEPES, pH 8.0, 75 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 0.5 mM DTT, 4 mM NaF, 12% glycerol) containing 1 μg pUC9 DNA (cut with HpaII), 10 μg yeast tRNA, 1 μg dNTPs and 15-50 μg of nuclear extract. 10 units of exonuclease III (Boehringer) were added and incubation was continued for 12-20 min. The reaction was stopped with 25 μl of 350 mM ammonium acetate, and 10 mM EDTA. The protected DNA was purified by phenol-chloroform extraction, precipitated with ethanol and analyzed on 6% denaturing acrylamide gels.

**RESULTS**

The 3' end of rat pre-rRNA maps 560-565 nt downstream of the 28S rRNA terminus (unpublished data). This site of 3' end formation of rDNA transcripts is located immediately upstream of a cluster of 18 bp conserved sequence elements which are present eight times in the 3' spacer between nucleotides +582 and 1767 (see Figures 1 and 3, ref. 5). Since the number, localization

---

**Figure 1**: Diagram of the 3' spacer of rat and mouse rDNA and the subclones used. The location of the 18 bp Sal I box elements is marked by boxes. The numbers indicate their position in nucleotides relative to the 3' end of 28S rRNA. The rat spacer regions that have been subcloned and fused to the mouse rDNA promoter are shown. The portion of the spacer that is contained in 45S pre-rRNA is indicated by a wavy line. The arrowhead marks the Sau3A restriction site at position +352.
Figure 2: Transcripts derived from rat rDNA minigenes
A) 100 ng of pMPRT1-6 template DNA in the circular form (lane 1) or after linearization with Eco RI (lane 2) were incubated in a 50 μl standard transcription assay (3). The transcripts terminated at the different Sal I box elements are designated T\textsubscript{4} to T\textsubscript{6}; RT marks readthrough transcripts.
B) Transcripts synthesized from circular templates pMPRT 7 (lane 2) and pMPRT 8 (lane 3), respectively. For control, transcriptions with the mouse minigene construct pPTBH/Eco RI (lane 1) have been performed as well (3). Transcripts terminated in front of the first or second termination site contained in pPTBH are labelled T\textsubscript{1} and T\textsubscript{2}.

and primary structure of these elements is very similar to the murine transcription termination signal -the Sal I box (AGGTGACCAG\textsubscript{AT\textsubscript{1}}T\textsubscript{1}NTCCG)- present in the 3' terminal rDNA spacer of mouse, they are likely to serve a similar function. To investigate the functional role of these sequence motifs, we fused different regions from the rat spacer to a mouse rDNA promoter fragment and tested these artificial minigene constructs in extracts from Ehrlich ascites cells for their ability to direct the synthesis of correctly terminated RNA molecules.

Three different mouse-rat hybrid gene constructs were used encompassing different regions from the rat spacer. pMPRT 1-6 contains the rat Sal I box elements T\textsubscript{4} to T\textsubscript{6}, pMPRT 7 and pMPRT 8 the elements T\textsubscript{7} and T\textsubscript{8}, respectively (Fig. 1). pMPRT 1-6 in the circular form yields five distinct RNA bands (Fig. 2A, lane 1) the lengths of which correspond to transcripts that have been initiated at the mouse transcription start site and terminated in front
Figure 3: Alignment of nucleotide sequences contained in the different rat boxes and summary of their biological activity, as determined in transcription and exonuclease III protection experiments. The nucleotide sequence of the 3' spacer from mouse and rat rDNA has been published (3, 5). The two domains corresponding to completely conserved bases in all mouse Sal I boxes are framed. The nucleotides exchanged in the rat signals are encircled. Transcription termination in the mouse system with 50-100% of maximal activity is marked by a +, between 25 and 50% of maximal activity by -, and below 10% by -. The binding of the factor to sites T₄, T₅ and T₆ has not been determined in exo III protection experiments (n.d.).

An alignment of the nucleotide sequences of the eight presumptive rat termination signals T₁-T₆ is shown in Fig. 3. Each element contains one to four base exchanges as compared to the highly conserved mouse sequence. Obviously
Figure 4: Effect of rat spacer DNA on transcription termination of mouse rDNA in vitro. 10 ng of the mouse minigene pPTBH was linearized with Eco RI and transcribed in vitro in the presence of 90 ng pUC9 (lane 1) or pRT1-6 (lane 2). In parallel, the same amount of competitor DNA pRT1-6 (lane 3) or pUC9 (lane 4) was preincubated for 15 min before addition of template DNA. Transcripts terminated at sites T₁ or T₂ and readthrough transcripts (RT) are marked.

certain base substitutions within the 18 bp sequence motif at nucleotides 1, 5, 10, 16, and 18 do not significantly impair function as revealed by the ability of sites T₁, T₂, T₄ and T₆ to direct transcription termination. The C to T transition at position 17 (T₇) reduced termination activity by more than 50%. Also natural mutations of nucleotides 9 (sites T₃ and T₇) and 15 (sites T₃ and T₈) strongly reduced or abolished termination. These findings suggest that the mutations differently affect the interaction of a specific DNA-binding protein with the Sal I box consensus sequence.

In order to demonstrate the binding of the mouse termination factor to the rat Sal I box sequence two different experimental approaches were used - template competition and exonuclease III protection experiments. Fig. 4 shows the effect of increasing amounts of rat spacer DNA contained in pRT1-6 on the formation of correctly terminated transcripts from the murine minigene construct pPTBH (3). Linear pPTBH directs the synthesis of three discrete RNA bands which reflect readthrough RNA (RT) or transcripts terminated at site T₁ or T₂, respectively. At low template concentrations in the presence of a 9-fold excess of unspecific competitor DNA (pUC9) the majority of transcripts is terminated at site T₁ (lane 1). In the presence of an excess of rat spacer DNA pRT1-6 a shift from terminated to readthrough transcripts is observed (lane 2). Preincubation of the reaction mixture with the specific (lane 3) or nonspecific competitor DNA (lane 4) before addition of the
Figure 5: Binding of the mouse termination factor to the rat Sal I box sequences. A 292 bp Eco RI-Sal I fragment from pRT12, which contains rat spacer rDNA sequences from +352 to +632 (see Fig.1) was 5' end labeled at the Eco RI site. After incubation with mouse extracts the samples were digested with exonuclease III as described in Material and Methods. Lane 1, untreated fragment, lane 2, exonuclease III protection assay performed in the presence of S-100 extract, lane 3, exonuclease III protection assay performed in the presence of 2/ul nuclear extract. M, size marker (pBR322/Hpa II).

Template yields the same quantitative distribution of RNA bands. The suppression of transcription termination from the mouse minigene by the rat spacer sequences indicates that the mouse termination factor has been specifically sequestered by the rat rDNA.

The specific binding of the mouse nuclear factor to the rat Sal I box elements was demonstrated directly by exonuclease III protection experiments. A 292 bp Eco RI-Sal I fragment derived from pRT12 was 5' labelled at the Eco RI site and was incubated with either S-100 or nuclear extracts. The assays with the S-100 extracts served as a negative control, because the amount of termination factor present in these cytoplasmic extracts is very low (unpublished results). As shown in Figure 5, in the presence of nuclear extracts a new 255 nt fragment was obtained after exonuclease III treatment (lane 3). The length of this exo III-resistant fragment maps the 3' boundary of the binding site of the mouse factor to position 607, i.e. a few nucleotides beyond the Sal I box sequence. Thus the binding region on the rat element corresponds to that previously determined on the mouse signal sequence (4).

The fact that the mouse factor recognizes the rat box enables us to determine the effect of the natural point mutations in the different rat elements on
their interaction with the Sal I box-binding protein. To quantitate the relative affinity of the binding protein to the various sequence elements exonuclease III protection experiments were carried out in the presence of competitor DNAs containing different regions from the rat spacer. For this, a 202 bp Eco RI-Hind III fragment from the mouse clone pUCT was 5' end-labelled at the Eco RI site. After incubation with nuclear extract and two different concentrations of competitor DNA the samples were treated with exonuclease III and analyzed by gel electrophoresis. As shown in Fig. 6 specific binding of the mouse nuclear factor generates a new 180 bp exonuclease III-resistant fragment the strength of which depends on the amount and type of competitor DNA used. Increasing amounts of both the mouse plasmid pUCT (lanes 4 and 5) or the rat plasmid pRT12 (lanes 6 and 7) similarly competed for factor binding, resulting in a decreased intensity of the 180 bp fragment. This finding indicates that the ability of the Sal I box-binding protein to interact with either the mouse or the rat termination signal is comparable. However, if the reaction was performed in the presence of pRT7 or pRT8 no competition was observed (lanes 8 - 10). The same result was obtain-
ed with a subclone containing site T₃ (data not shown). These results are in
accord with the transcription data (see above) which revealed no tran-
scription termination from minigene constructs containing sites T₃, T₇ or T₈
indicating that these signals are biologically inactive.

DISCUSSION
Previously we and others have shown that the RNA polymerase I (pol I) tran-
scription machinery exerts a marked preference for homologous rDNA. Both in
vitro transcription experiments and transfection studies revealed the in-
volvement of species- or order-specific factors in pre-rRNA synthesis
(9 - 11). This evolutionary change of transcription factors has been accom-
panied by a molecular co-evolution of signal sequences that direct faithful
transcription initiation by pol I (12, 13). A similar co-evolution appears
to have occurred between the sequences and factors that direct correct 3' end
formation of pre-rRNA molecules. In mouse we have shown that a repeated 18 bp
signal in the 3' spacer - the Sal I box - functions as transcription termi-
nator, mediating stop of the elongation reaction and release of nascent RNA
chains. This sequence motif is the primary target for a specific DNA binding
protein, and the interaction of this protein with the Sal I box sequence
motif has been shown to be a prerequisite for transcription termination (3,
4).

The same signal sequence is found several times at a similar location in the
spacer of rat rDNA (5). Moreover, a cluster of conserved 10 bp sequence
elements with an 8 bp perfect homology to the palindromic part of the mouse
box is found in the human 3' terminal NTS region downstream of the 28S rRNA
coding region (14, 15). The human sequence is functionally analogous to the
mouse pol I terminator. We have shown that it mediates binding of a nuclear
protein which functions as a termination factor (16). The murine signal
sequence is recognized by the human factor but not vice versa. Thus the
change of the termination signal sequence apparently has been accompanied by
a molecular co-evolution of the gene(s) coding for the termination factor.
The molecular mechanisms which may be involved in spreading new information
through a multigene family have in detail been discussed before (17, 18).
A sequence comparison of rDNA transcription units of various lower eu-
karyotes reveals no significant sequence homology among the sites of 3' end
formation (19, 20). Most interestingly, alternative strategies for the ge-
genation of correct 3' ends have been developed in different organisms.
Whereas in yeast pre-rRNA synthesis is stopped at a termination site 210 bp
downstream of the 25S rRNA end (19), in *Xenopus laevis* and *Drosophila melanogaster* the spacer is almost entirely traversed by RNA polymerase I. The 3' end of both 28S RNA and 40S pre-rRNA is formed by a processing event. The primary transcript is very unstable and appears to terminate upstream of the initiation site (1, 2, 21).

Such an upstream terminator has also been identified in the vicinity of the mouse rDNA transcription initiation site (22, 23) and its participation in modulating the initiation efficiency has been demonstrated (22). An alignment of promoter sequences of different mammals, i.e. human, chimp, rhesus, mouse, rat and rabbit (24), reveal the conservation of a Sal I box-like sequence present in all species about 170 bp upstream of the initiation site. This remarkable conservation of both the position and nucleotide sequence of this putative upstream terminator suggests an important role in transcription initiation.

In the various species examined this upstream terminator sequence contains a few point mutations as compared to the mouse Sal I box. Similarly, the repeated Sal I box-elements present in the 3' terminal spacer of rat rDNA differ by one to three base exchanges from the highly conserved mouse sequence. Obviously, the mechanisms that maintain homogeneity both amongst the Sal I boxes in a given repeat and, of course, between all or most rDNA repeats have ceased to operate to a certain extent in rat leaving each Sal I box to accumulate mutations independently. Such mechanisms would be unequal crossing-over or slippage.

In this paper we have addressed two questions: (I) Do mouse and rat use the same signals and factors to accomplish termination of rDNA transcription, and (II) how do individual base exchanges affect the termination process? We have used a heterologous system and artificial mouse - rat minigene constructs and demonstrated that the rat signal can functionally replace the mouse sequence. Termination evidently requires a trans-acting factor which binds to the Sal I box sequence and can be titrated with both mouse or rat DNA. Under the experimental conditions used termination functions only to a limited extent (see Fig. 2A). This is due to the fact that the factor is present in the extracts at limited amounts. Thus the chance that a given template molecule with several Sal I box elements will have bound both all the proteins required for initiation and more than one termination factor is relatively low. Therefore the termination reaction appears to be somewhat leaky. Complementation of the extracts with partially purified mouse factor eliminates read-through transcription and results in a preferential ter-
mination at the first box (I. Bartsch, unpublished results).

On the other hand, the fact that there is no excess of the Sal I box binding protein in the extracts enables us to detect differences in the relative binding strength of individual rat Sal I box elements. We have found that the mouse factor interacts efficiently with most of the rat boxes although each of them contains one or more point mutations. This result apparently contradicts previous findings where we have introduced base substitutions into the mouse box at positions 3 and 4, or 15, by site-directed mutagenesis. In these mutants factor binding and transcription termination was erased or strongly impaired (4). This susceptibility of the termination reaction to base alterations in the termination signal suggested that changes of the Sal I box sequence result in loss or reduction of biological function.

The results presented in this communication, however, clearly show that certain base exchanges are tolerated without affecting factor recognition and termination activity. This is especially obvious for the first and last nucleotide of the 18 bp sequence. There both a guanine, adenine or cytosine is present in functionally active terminator elements. Also the C to A transversion at nucleotide 5 (box T5) or an A to G transition at position 10 (box T6) does not impair function. On the other hand, the elements where the C residue at position 9 is converted into a thymidine are incompetent for factor binding and termination (boxes T3 and T7). Since, however, these elements contain two more base alterations in the distal part of the consensus sequence a valid quantitation of the role of single nucleotides within the termination signal sequence is hard to assess.

The highest mutation frequency is observed in the distal part of the Sal I box - the TCCG motif. Obviously a G to C transversion at position 18 does not affect the biological function (sites T1, T4 and T6), whereas the single C to T transition at position 17 reduces termination activity approximately 2-3 fold (Fig. 3). A similar inhibition was observed when the thymidine of the TCCG motif was converted into a cytidine residue by site-directed mutagenesis (4).

The functional importance of a few bases within a larger region that is recognized by a specific transcription factor has already been demonstrated for the rDNA promoter. A series of point mutations has been introduced into the proximal domain of the mouse rDNA promoter. Analysis of the transcriptional activity and initiation factor binding showed that out of the 22 mutants examined, only changes at positions -1, -7, and between -16 and -21 affected biological activity although a region of at least 39 bases was
required for minimal promoter function (25). The elucidation of the role of single nucleotides within specific transcriptional signal sequences will require the isolation of the trans-acting factors which recognize these sequences. Studies on the exact residue - base contacts between proteins and DNA as well as between different proteins will finally unravel the molecular mechanisms involved in the selectivity and regulation of gene expression.

**ACKNOWLEDGEMENTS**

We thank Ingrid Bartsch for critically reading and E. Zimmer for typing the manuscript. We would also like to thank FEBS for providing M. Kermekchiev with a fellowship enabling him to work in Würzburg. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Sonderforschungsbe-reich 165) and the Fond der Chemischen Industrie.

*Permanent address: Institute of Cell Biology and Morphology, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

**REFERENCES**