The two main rDNA size classes of Ascaris lumbricoides: comparison of transcription termination and spacer organization

Eliane Müller+, Heinz Neuhaus§, Heinz Tobler and Fritz Müller*
Institute of Zoology, University of Fribourg, Pérolles, CH-1700 Fribourg, Switzerland

Received April 8, 1992; Revised and Accepted May 28, 1992
GenBank accession nos M74584 and M74585

ABSTRACT
Structural and functional analyses were carried out to compare transcription termination and intergenic spacer organization between the two heterogeneous ribosomal DNA size classes of Ascaris lumbricoides. By performing mung bean nuclease mapping in vivo, we localized the 3' end of the mature 26S rRNA to the same position in both forms. This site coincides with the in vivo and in vitro transcription termination site of the 40S-precursor rRNA from both rDNA size classes. We demonstrate that the 3' boundary of sequences necessary for faithful transcription termination extends 35 bp beyond the end of the 26S rRNA gene and that these sequences are perfectly conserved within the two size classes. The intergenic spacers of both rDNA size classes have been completely sequenced; they are 2410 bp and 1937 bp long and show a sequence homology of 92.8%. There is no evidence for the presence of reiterated termination sites or promoters within the relatively short intergenic spacers of the two rRNA gene classes.

INTRODUCTION
In most eukaryotes the ribosomal RNA genes are arranged in clusters of several hundred tandemly repeated transcription units. Members of such multigene families undergo 'concerted evolution', implicating chromosomal rearrangements for the propagation and fixation of novel variants (see ref. 1, for review). The consequence of this phenomenon is an almost complete homogeneity within the ribosomal gene family of a given species.

The rDNA family of Ascaris lumbricoides is not homogeneous but rather split up in two distinct subfamilies, which are easily distinguished by a 486 bp long insertion in the intergenic spacer of the long form (2, 3). The quantitative distribution of these two families is about 10:1 in a pool of animals, but can vary within individuals (4). Although both size classes are located on the same autosome, they are not intermingled with each other (4, 2). Based on Southern blot experiments and sequence analyses, we present here evidence that 'concerted evolution' occurs within but not between the two rDNA subfamilies. Because both forms are transcribed (3), we expect that the transcription regulatory domains are conserved between the two rDNA size classes. In this paper we mainly focus our attention to a comparison of the regulation of transcription termination and the organization of the intergenic spacer.

Recently we have reported that the longer rDNA form of Ascaris mediates transcription termination in vivo and in vitro at the 3' end of the 26S rRNA gene (5). This is a very unusual feature since all other organisms investigated so far use this site as a processing site, whereas the terminator is localized further downstream in the intergenic spacer (6). Here we show that also the short organizational rDNA form terminates transcription at the 3' end of the 26S rRNA gene. The 3' boundary of the sequences required for accurate in vitro transcription termination lies within a short region which is conserved between both rDNA size classes and is located immediately downstream of the end of the 26S rRNA gene. Moreover, we have completed the sequence of the intergenic spacer of both rDNA forms and compared it to those of other species.

MATERIALS AND METHODS
DNA sequencing
We followed the procedure of Sanger et al. (7) for the establishment of the nucleotide sequences of the cloned representatives of the two rDNA size classes of Ascaris lumbricoides.

Plasmid constructions
The original clones used for all subcloning experiments are pAlr8 and pAlr13 (4), containing a single repeating unit of either the long or the short size class of Ascaris rDNA. The miniplasmids tested in the in vitro assay were constructed by inserting different restriction fragments into the poly linker of clone HHi661 (5), downstream of a promoter element. The 3' deleted Alul-Alul fragments described in Fig. 4A and 5A were created by progressive digestion with exonuclease III following the method

* To whom correspondence should be addressed

Present address: + Centre Hospitalier Universitaire, Faculté de médecine, Service d'immunologie, Université de Sherbrooke, Sherbrooke J1H 5N4, Canada and § Baxter and Dade, CH-3186 Düdingen, Switzerland
of Henikoff (8). The truncated fragments were sequenced using the dideoxy method for double strand sequencing suggested by the supplier of the Sequenase™ Vers. 2.0 (Stratagene, San Diego).

**Mung bean nuclease mapping**

Total RNA from oogonia, oocytes and larvae of *A. lumbricoides* was isolated following the method of Pilgrim (9). The larval RNA was fractionated on a sucrose gradient and the 40S-precursor rRNA and mature 26S rRNA were collected separately. The DNA probe was radioactively labelled at the 3′ end by α-32P-dGTP (800 Ci/mmol; Amersham). Roughly 15,000 cpm of the labelled *Mspl/Xbal* fragment (cf. Fig. 3A) were mixed with 5–10 μg of the corresponding RNA. The mung bean nuclease mapping was carried out with 100 μl/ml of enzyme for 30 min at 37°C (10). The sequencing reactions for the sizing of the digestion products were performed following the method of Maxam and Gilbert (11).

**Southern blot hybridization**

For Southern transfer and hybridization under high stringency conditions we followed essentially the standard techniques (10).

**In vitro transcription assay**

Whole-cell extracts were prepared from oogonia and used essentially under the same transcription conditions as reported previously (3). The total volume of each assay was 12.5 μl and contained 0.3–0.5 μg of plasmid DNA.

## RESULTS

Nucleotide sequence comparison between the 26S rRNA coding and the adjacent intergenic spacer regions of the two ribosomal size classes

We have determined the nucleotide sequence of an rDNA region which comprises roughly 500 bp of the 3′ end of the 26S rRNA gene and 1000 bp of the adjacent intergenic spacer of the two main rDNA size classes present in *A. lumbricoides* (GenBank accession nos M74584 and M74585). They are almost identical within the sequenced part of the 26S rRNA coding region and within the first 40 bp of the adjacent intergenic spacer (Fig. 1B). Beyond this site (position +41 in the spacer), a deletion of 13 bp leads to a more diverged domain of the rDNA repeats. The rest of the spacer differs by several deletions, insertions and base substitutions. Our analysis demonstrates that the 26S rRNA genes of the two rDNA size classes, as far as sequenced, show a homology of 98.2%, which drops to 92.8% in the spacer regions. The nucleotide sequences established in this study, together with those reported in an earlier publication (3), encompass the complete spacer, as well as some part of the adjacent coding regions for both rDNA forms (Fig. 1A). The intergenic spacers of the long and short rDNA size classes are 2410 bp and 1937 bp long, respectively.

Sequence microheterogeneities are representative for the two rDNA size classes

The main difference between the two ribosomal size classes of *Ascaris* resides in a 486 bp long insertion in the intergenic spacer.
of the long form (cf. Fig. 1A). Hybridization of XbaI/BamHI digested genomic DNA with the 1.4 kb long EcoRI fragment as probe (see Fig. 2B), lights up the 3100 bp long XbaI-BamHI fragment (I) together with the 2390 bp long XbaI spacer fragment of the long form (II, lane 8 and G), and the corresponding 1904 bp long fragment of the shorter form (II', lane 13 and G). The relative intensity of the hybridization signals agrees with the earlier described quantitative genomic distribution of the long and the short form of roughly 10:1 (4). An additional band beyond fragment I corresponds to an rDNA form carrying a type I insertion in the 26S coding region (12). Beside this major structural difference, several microheterogeneities are found in the sequences between the clones pAlr8 and pAlr13 (Ref. 3); there is, e.g., an additional MboII site in the spacer of the long form, which is missing in the short clone because of a 13 bp deletion (cf. Fig. 1B). MboII digestion of the 1.4 kb long EcoRI fragment of the long form containing this region (Fig. 2C) yields a 500 bp and a 900 bp long fragment (III, IV), but does not cleave the corresponding EcoRI fragment of the short form (III'). In order to check whether this restriction polymorphism is representative for all members of a particular rDNA size class, we performed a Southern hybridization experiment with EcoRI and MboII digested genomic DNA. Fig. 2C shows that the expected restriction fragments are indeed present, and the ratio of their hybridization signals of about 10:1 corresponds again to the overall genomic distribution of the long and the short rDNA form (Fig. 2B). A few other restriction site polymorphisms have been analyzed in the same way, leading to identical results (data not shown). The quantitative distribution of the restriction polymorphism tested in all investigated cases always matched the overall distribution of the two rDNA size classes, thus providing strong evidence that the microheterogeneities as observed between the two sequences of the cloned rDNA forms are representative for their respective size classes.

**Fine mapping of the 3' end of the rRNA transcripts in vivo**

We have shown previously by S1-mapping experiments that both rDNA size classes are transcribed in vivo (3). In order to precisely determine the 3' end of the corresponding 26S rRNA, we performed a mung bean nuclease mapping experiment with RNAs from different developmental stages. Thus, the end-labelled MspI-XbaI fragment of clone pAlr8 was hybridized to 40S-precurso rRNA and 26S rRNA isolated from larvae, as well as to total RNA from oocytes and oogonia (Fig. 3A and B). The mung bean nuclease digestion products were then run along with the sequencing ladder of the same MspI-XbaI fragment. The autoradiograph demonstrates that the larval 40S-precurso rRNA and 26S rRNA, and the rRNAs from oocytes and oogonia, share the same 3' end in vivo. Nuclease digested fragments do not co-migrate with the corresponding fragments resulting from the chemical sequence reactions, but 1½ nucleotides slower (13). Therefore, the major band in all DNA species maps to a guanosine (indicated by position -1), which is located 129 bp downstream of the labelled MspI site. Several additional minor bands may result from either an inaccurate digestion of the nuclease, an ambiguity in transcription termination, or a rapid transcription-coupled processing event that removes a few nucleotides. The latter possibility, however, is an unlikely explanation, because there is no difference in the banding pattern between 40S precursor and 26S rRNA. The absence of a hybridization signal in the upper portion of the gel (not shown), corresponding to the intact probe, confirms that no cellular rRNA

---

**Figure 2. Genomic organization and distribution of the spacer region of the long and the short rDNA size classes in Ascaris.** (A) Restriction map of the Ascaris rDNA for the enzymes applied in this experiment. The 1.4 kb long EcoRI fragment was used as hybridization probe for the Southern blots in (B) and (C). The insertion (INS.) in the long rDNA form is indicated by a box. (B) Restriction map for the enzymes BamHI and XbaI of the spacer regions of clones pAlr8 and pAlr13, representing the two rDNA size classes. The Southern blot contains BamHI/XbaI digested pAlr8 (8) and pAlr13 (13) DNA, as well as total larval DNA (G), hybridized with the 1.4 kb long EcoRI fragment as probe. (C) Restriction map of the 1.4 kb long EcoRI spacer fragment, showing a restriction length polymorphism for MboII between both forms which is created by a 13 bp long deletion in the spacer of the short form (cf. Fig. 1B). At the right side there is the autoradiograph of a Southern blot containing EcoRI/MboII digested pAlr8 and pAlr13 DNA, as well as total larval DNA, hybridized with the 1.4 kb long EcoRI fragment as probe.
Figure 3. Precise mapping of the 3' end of 26S and 40S-precursor rRNA. (A) Schematic representation of a part of the long Ascaris ribDNA clone pAlr8. The MspI-XbaI fragment covering the end of the 26S rRNA gene was labelled at its 3' end with Klenow DNA polymerase. The labelled fragment was used for chemical sequencing (12) and as probe for the precise in vitro mapping of the 3' end of the rRNA transcripts. (B) The labelled MspI-XbaI fragment was hybridized to 40S-precursor rRNA, to 26S rRNA of larvae, as well as to total RNA from oocytes and oogonia. The hybrids were digested with mung bean nuclease (lOOU/ml), denatured and run along with the corresponding sequencing ladder on a 6% polyacrylamide gel. In rRNA species, the main band maps to the guanosine at the 3' end of the mature 26S rRNA, designated as position -1 in Fig. 3A.

Sequence requirement for transcription termination in vitro

In vitro transcription termination in the long form of the Ascaris ribosomal genes occurs at the end of the 26S rRNA gene (5). To analyze transcription termination in the short form, we constructed the miniplasmid HpAA(13). This plasmid carries a polymerase I promoter upstream of the 439 bp AluI fragment, covering the 3' end region of the 26S rRNA gene of pAlr13 (cf. Fig. 4A). HpAA(13) was incubated under standard conditions in our homologous in vitro transcription system (3). As reference we used the plasmid HpAA(8), containing the corresponding AluI fragment of the long form, known to be able to mediate species longer than about 130 nt are present. Since the nuclease protected sequences between the MspI site and the end of the 26S rRNA coding region are identical between both clones, this result is also representative for the transcripts of the short form.

Figure 4. In vitro 3' end formation of the transcripts of the short and the long rDNA form. (A) Plasmid HpAA(13) contains the 439 bp long AluI-AluI fragment with the 3' end of the 26S rRNA gene of the short form (pAlr13) downstream of a polymerase I promoter. This fragment covers 190 bp of the end of the 26S rRNA gene together with 249 bp of the bordering spacer sequences. Plasmid HpAA(8) carries the corresponding 437 bp long AluI-AluI fragment of the long rDNA form. Upon exonuclease III digestion of the two plasmids, the following truncated clones were generated: Hp(13)A +93, a digestion product of HpAA(13), a digestion product of HpAA(13), containing 93 bp of spacer sequences downstream of the termination site. Hp(8)A +35 and Hp(8)A +15 are derived from the long form and include 35 bp or 15 bp, respectively, of the intergenic spacer. (B) In vitro transcription of the above indicated clones, either digested with HindIII or in their supercoiled form (sc). All transcription products were run on a 4% polyacrylamide gel. The 307 nt long termination product is indicated by a black arrow, whereas the different run-off transcripts are marked by white arrows. As control, the same amount of extract was incubated under standard transcription conditions but without adding any further DNA. Read through transcripts are designated with R.T. Lane M shows HindIII digested pBR322 as a size marker.
transcription termination (5). Upon digestion with HindIII, both templates generate a 307 nt long rRNA transcript, extending to the 3' end of mature 26S rRNA, as well as a 571 (573) nt long run-off transcript (Fig. 4B, lanes HpAA(13), HpAA(8)). Thus, our results clearly prove that the Alul fragment of the short form, although diverging in its nucleotide sequence from the long form in the intergenic spacer downstream of position +40, is able to terminate transcription in vitro as efficiently as the corresponding fragment of the long form.

In order to narrow down the 3' boundaries of the sequences required for transcription termination, we have introduced various deletions at the 3' end of the spacer fragments in the clones HpAA(8) and HpAA(13) by progressive exonuclease III digestions. The extent of the deletions were determined by sequencing. Plasmid Hp(13) △ +93, a deleted clone of the short rDNA form, contains the terminal 190 bp of the 26S rRNA gene and 93 bp of the adjacent intergenic spacer (Fig. 4A). Hp(13) △ +93, linearized with HindIII, forms a correct 3' end (band at 307 nt), as well as a run-off product extending to the HindIII site of the polylinker (Fig. 4B). The ratio of terminated to run-off rRNA is the same as for the non-deleted plasmid HpAA(13). We then tested the clone Hp(8) △ +35 which extends to position +35 in the spacer, covering the conserved sequences next to the termination site (cf. Fig. 1B). The linearized clone mediates termination in an in vitro extract, but with a significantly lower efficiency than the control clone HpAA(8), as shown by a very weak band at 307 nt length. When incubating the same amount of plasmid in the supercoiled form, termination occurs at normal rate (cf. Fig. 4B lane Hp(8) △ +35 sc.). The shortest clone Hp(8) △ +15, however, if transcribed in linearized or supercoiled form, is no longer able to terminate transcription and forms only run-off rRNAs or read through transcripts. From these results we conclude that the 3' boundary of the sequences, necessary for correct transcription termination in both forms, is located close to the 3' end of the conserved spacer region, which extends 40 bp downstream of the transcription termination site. These findings, together with our earlier published data (5), allow us to localize the region essential for 3' end formation and transcription termination within the 105 bp long region located between position -70 in the 26S rRNA gene and position +35 in the intergenic spacer.

Clone Hp(8) △ +35 was shown to terminate transcription more efficiently in the supercoiled than in the linearized form (see Fig. 4B), suggesting that termination efficiency depends on the length of DNA sequences present at the 3' site of the termination signal. In order to test this hypothesis, we have compared the in vitro transcription termination efficiency of plasmid Hp(8) △ +35, either in its supercoiled form or digested with HindIII or PvuII. The restriction enzyme HindIII cuts the plasmid at the 3' end of the 20 bp long polylinker, located downstream of the 35 bp long rDNA spacer fragment, whereas PvuII digestion leaves the entire polylinker together with 91 bp of adjacent vector sequences (Fig. 5A). Incubation of equal amounts of plasmid DNA demonstrates that the efficiency of in vitro termination is higher for the supercoiled form and the PvuII linearized plasmid than for the HindIII digested DNA (Fig. 5B). We therefore conclude that the length of DNA sequences present downstream of the termination site (even if non-specific), is important for the efficiency of transcription termination in vitro.

Overall structure of the intergenic spacer

The primary structure of the intergenic spacers of both rDNA clones has been analyzed by a computer search for subrepeats, reiterated terminator and promoter elements, as well as for open reading frames. We also tested the spacers for homologies to other genes.

Apart from the long direct repeats flanking the 486 bp heterology in the spacer of the long form (3), no significant repeated sequence structures are detected. Even by carefully screening the entire spacer with the computer aided Pustell's matrix comparison program (14), we were unable to detect any indication for the presence of repeated promoter regions and termination sites. The absence of multiple termination sites was already demonstrated by a functional analysis (5). Furthermore, no transcriptional activity was detected in the 5' part of the intergenic spacer by using a nuclear run-on transcription assay in vitro (5). Screening of the coding and the non-coding strand revealed a single ORF in the long form, which encodes a potential polypeptide of 126 amino acids length. No homology to a known protein, however, could be found. A search for sequence
homologies to other genes and DNA elements which have been reported to be present within the spacer of ribosomal genes of various species, including 5S rRNA genes (15), tRNA genes (16, 17), ARS (autonomously replicating sequences; 18) and the enhancer sequence for ribosomal gene transcription in yeast (19), yielded no positive result.

DISCUSSION

Southern blot analyses demonstrated that restriction site polymorphism between cloned representatives of the two Ascaris rDNA size classes occurs in the genomic rRNA genes in the same quantitative ratio, namely about 10:1. It therefore seems to be likely that the nucleotide sequences of the two cloned members are representative for the majority of the rRNA genes within the corresponding size class. This hypothesis is further supported by two other observations: i) an additional restriction site within the 26S rRNA gene is present in four different clones of the long size class, but absent in four clones of the short size class (4); ii) a dT5 cluster in the external transcribed spacer of the short clone is replaced by a dT3 cluster in the long form. The rRNA transcripts carrying an U5 cluster at this position were shown to support transcription termination in yeast (19), but only very few nucleotide changes are observed within the coding regions of the two ribosomal gene classes of the chromosomal rRNA genes (20), which are localized on a single chromosomal locus (21). Accumulation of a large number of nucleotide changes between the two rDNA size classes of A. lumbricoides strongly indicates that both clusters are physically separated and evolve independently. Therefore, although localized on the same chromosome (2), they have to reside in two different chromosomal loci. A case for independent evolution of members of a multigene family is reported for the 5S rRNA genes in Neurospora crassa, which are distributed on different chromosomal loci (22).

Whereas the primary structure of the spacers in the two size classes of the Ascaris rRNA genes diverged relatively largely, only very few nucleotide changes are observed within the coding regions as far as sequenced (this paper and Ref. 3). Obviously, the coding regions were maintained homogeneous in evolution because of important functional constraints. Regulatory domains seem to be subject to the same evolutionary pressure since the minimal promoter required for transcription initiation is completely identical (3). Furthermore, we show here that the conserved sequences between the two ribosomal gene classes extend about 40 bp into the adjacent intergenic spacer and that this region is required for correct 3' end formation and termination of the rRNA transcript in vivo. Mung bean nuclease mapping (this paper) and very sensitive RNAse mapping experiments (5) demonstrate that transcription termination occurs within this short conserved region at the junction of the 26S rRNA gene and the intergenic spacer. Short sequence boxes downstream of the termination sites were reported to be sufficient to mediate transcription termination in mammals and Xenopus (6). In contrast to vertebrates, however, transcription termination in Ascaris requires additional regulatory sequences upstream to the end of transcription. Their 5' boundary is located between position —70 and —20 (this paper and Ref. 5). The complete region sufficient to support transcription termination in vitro, therefore, is located between nucleotides —70 and +35. Within this region no sequence homologies were found between the Ascaris termination signals and those reported for other species (6). It is likely that termination of transcription in Ascaris is also brought along, like in vertebrates, by specific DNA binding proteins (23). This hypothesis is supported by the higher efficiency of in vitro termination on templates with extended, non-specific DNA stretches present downstream of the 3' boundary of the termination signal. Alternatively this non-specific DNA stretch may just be required by the Poll itself to terminate transcription efficiently. To answer these most interesting questions, bandshift and footprint experiments in parallel with site directed mutagenesis have to be done.

No repeated termination sequences are present within the Ascaris intergenic rDNA spacer. This is in good agreement with the results of our functional analyses, confirming the absence of additional terminators in the spacer downstream of the authentic termination site (5). Multiple termination sites, however, are found in the spacer and in front of the next initiation site of the vertebrate rRNA genes (6), for review, the latter preventing polymerases reading into the initiation region (24). Such fail safe terminators seem not to be necessary for the ribosomal genes in Ascaris, since there is no evidence for polymerase I spacer transcription (5). Read though transcripts may either not exist at all, thus presupposing that the terminator at the end of the 26S gene is very efficient, or be stopped randomly at many non-specific sites within the intergenic spacer. Besides multiple termination sites, vertebrate and Drosophila rDNA spacers contain many repeated elements, which exert important regulatory functions, including promoter and enhancer activity (25, for review). No comparable repeats are present in the relatively short spacer of Ascaris, which in this respect is similar to the structural arrangement of the ribosomal spacers in Caenorhabditis elegans (26) and S. cerevisiae (15). The spacer of the yeast rDNA, however, contains a single enhancer element, located about 2 kb upstream of the promoter (19). Based on structural homologies, there is no evidence for the existence of such an element in the rDNA of Ascaris. Functional tests will be needed to clarify whether analogous elements do exist.

Altogether, our structural and functional data indicate a different regulatory principle for rRNA gene transcription in Ascaris than in vertebrates. The ribosomal spacers of Caenorhabditis and yeast, however, are organized in a similar way than those of Ascaris, thus indicating an analogous regulatory mechanism. Drosophila, on the other hand, in which no termination activity for polymerase I transcription was detected (27), stands apart from lower eukaryotes, as well as from vertebrates. Transcription initiation of the ribosomal genes between different species was shown to share little common features (28). The present results suggest that the mode of transcription regulation, and especially termination of transcription, diverged even more during evolution.

ACKNOWLEDGMENTS

We would like to thank Mrs. Christine Stempfel and Mr. Hubert Gachoud for excellent technical assistance. We are grateful to all colleagues of our laboratory, in particular to Dr. Adrian Etter, for helpful discussions. This work was supported by grants No
3.113-0.85 and 3.141-0.88 from the Swiss National Science Foundation.

REFERENCES