Volume 11 Number 6 1983 Nucleic Acids Research

Nucleotide sequence of the transcriptional initiation region of *Dictyostelium discoideum* rRNA gene and comparison of the initiation regions of three lower eukaryotes' genes

Yoshiko Hoshikawa, Yoichi Iida* and Masaki Iwabuchi

Department of Botany, and *Department of Chemistry II, Faculty of Science, Hokkaido University, Sapporo 060, Japan

Received 17 January 1983; Accepted 18 February 1983

ABSTRACT

The 5' end of the rRNA precursor of *D. discoideum* was mapped on a cloned rDNA by S1 nuclease protection mapping, and the sequence of about 1240 nucleotides surrounding the transcriptional initiation site of the rRNA gene has been determined. Repeated sequences consisting of 16 nucleotides appeared in the region upstream from the initiation point. Comparison of the nucleotide sequences around the initiation site of rRNA genes in three lower eukaryotes, *D. discoideum*, *Saccharomyces cerevisiae* and *Tetrahymena pyriformis*, indicated that there was little similarity in the nontranscribed spacer regions, but in the transcribed spacer regions near the initiation point, very similar sequences consisting of 9 nucleotides were found.

INTRODUCTION

Recently, cell-free systems for transcription of rRNA genes by RNA polymerase I have been developed with crude cell extracts and cloned rDNAs (1-6). Several studies have indicated that cloned rDNA could faithfully be transcribed only with the homologous transcription system (5,6). To understand the species-specificity for the *in vitro* transcription of rRNA genes by RNA polymerase I, much data are required on the primary structure of the region surrounding the transcriptional initiation site of rRNA genes of various eukaryotes.

D. discoideum has approximately 180 copies of rDNA per haploid genome and rDNA exists as linear palindromic dimers without covalent attachment to chromosomal DNA, as in *T. pyriformis* and *Physarum polysephalum* (7). It has been suggested that the primary transcript of the *D. discoideum* rRNA gene is the 36S molecule, which is finally processed into mature 17S, 5.8S and 26S rRNAs (8,9).

In this study, we determined the nucleotide sequence of the

region surrounding the transcriptional initiation site of the *D. discoideum* rRNA gene cloned by Cockburn *et al.* (10) and also mapped the 5' end of the rRNA precursor. When the nucleotide sequence around the presumed transcriptional initiation point of the *D. discoideum* rRNA gene was compared with those of other eukaryotes, *S. cerevisiae* (11) and *T. pyriformis* (12,13), it was revealed that very similar sequences consisting of 9 nucleotides exist in the region near the initiation point, but they are located downstream from the initiation point.

MATERIALS AND METHODS

DNA preparation and DNA sequencing

Recombinant plasmid pDd 507 containing a part of *D. discoideum* rDNA, which was generated from randomly sheared nuclear DNA inserted into pMB9 by the poly(dA)-poly(dT) tailing method (10), was provided by R. A. Firtel. The recombinant plasmid was prepared by the cleared lysate method (14) and digested with restriction endonucleases under recommended conditions, followed by electrophoresis on 0.8 - 1.0 % agarose or 5 - 10 % polyacrylamide gels. DNA fragments were extracted from excised gel pieces and labeled at the 5' ends with [γ-^{32}P]-ATP and T4 polynucleotide kinase after alkaline phosphatase treatment as described by Maxam and Gilbert (15). The cleavage sites of restriction endonucleases on the rDNA fragment were mapped by partial digestion of the end-labeled DNA fragment as described by Smith and Birnstiel (16) and confirmed by single and double digestions of the same DNA fragment. DNA sequencing was performed as described by Maxam and Gilbert (15).

S1 nuclease protection mapping

To map the 5' end of the rRNA precursor and mature 17S rRNA on rDNA, S1 nuclease protection mapping was carried out as described by Berk and Sharp (17). The coding strand of the DNA fragment labeled at the 5' end with ^{32}P was hybridized either with 400 µg of whole nuclear RNA or 30 µg of 17S rRNA in 0.1 ml of 80 % formamide, 200 mM NaCl and 20 mM sodium citrate (pH 7.2) at 50^{o}C for 36 hr. After hybridization, the reaction mixture was diluted 1:10 with cold 50 mM sodium acetate (pH 4.5) containing 250 mM NaCl and 0.1 mM $ZnSO_4$, and then S1 nuclease

was added, followed by incubation at 45°C or 37°C for 30 min. The DNA-RNA hybrid protected from the S1 nuclease digestion was precipitated with ethanol, denatured and electrophoresed either on 1.5 % alkaline agarose gel (18) or 10 % sequencing gel (15).

Preparation of whole nuclear RNA and 17S rRNA

Whole nuclear RNA of *D. discoideum* A3 was extracted from isolated nuclei as described previously (19) and precipitated at least 3 times in 2 M LiCl, at 4°C for 24 hr after the first ethanol precipitation to remove nuclear DNA. Cytoplasmic 17S-rRNA was extracted from partially purified ribosomes of *D. discoideum* NC-4 and purified by two cycles of sucrose gradient centrifugation.

Enzymes and radioisotope

Restriction endonucleases were obtained from Takara Shuzo, Bethesda Research Laboratories Inc. and Boehringer Mannheim; bacterial alkaline phosphatase was from Bethesda Research Laboratories Inc.; T4 polynucleotide kinase from Takara Shuzo; and S1 nuclease from Boehringer Mannheim. [γ-^{32}P]ATP was purchased from Amersham.

RESULTS AND DISCUSSION

Nucleotide sequence of the roughly determined transcriptional initiation region

Although the restriction map of recombinant plasmid pDd 507 for Eco RI, Hind III, Pst I and Sal I has been reported by Cockburn *et al.* (10) and the coding regions of 17S, 5.8S and 26S rRNAs were roughly determined by Frankel *et al.* (20), there is no information about the nucleotide sequence of the region surrounding the transcriptional initiation site of the *D. discoideum* rRNA gene. So we made a more detailed restriction map of the 4.5 kb Eco RI/Sal I fragment on which both the 5' ends of the primary transcript of the rRNA gene and mature 17S rRNA are supposed to be located. Fig. 1 is a detailed restriction map of the Eco RI/Sal I fragment. Using this fragment we roughly determined the transcriptional initiation site of the rRNA gene. The end-labeled coding strand of the Eco RI/Sal I fragment was hybridized either with whole nuclear RNA or 17S rRNA and then digested with S1 nuclease. The size of the DNA

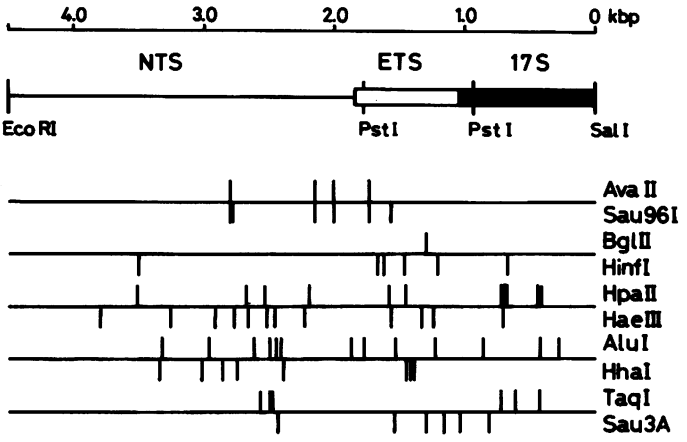

Fig. 1. Detailed restriction map of the 4.5 kb Eco RI/Sal I fragment. The open and closed boxes indicate the transcribed spacer region and the coding region of 17S rRNA, respectively.

fragments protected from S1 nuclease digestion was electrophoretically measured. There was a protected fragment 1.05 kb long when 17S rRNA was hybridized with the DNA probe (Fig. 2, lane 3). In the contrary, when the whole nuclear RNA was hybridized with the same DNA probe, the protection bands of 1.85 and 1.05 kb fragments were observed (Fig. 2, lane 4). These results indicate that the 5' end of 17S rRNA is located 1.05 kb upstream from the Sal I site and that the 5' end of the rRNA precursor in whole nuclear RNA is located a further 0.8 kb upstream from that of the 17S rRNA coding region. Since no band of DNA fragment larger than 1.85 kb was detected in the S1-mapping analysis of whole nuclear RNA (Fig. 2, lane 4), the rRNA precursor detected here is probably the primary transcript of the rRNA gene.

Then we determined the nucleotide sequence of the region surrounding the presumed transcriptional initiation site roughly mapped on rDNA by the method of Maxam and Gilbert (15). The restriction map and the sequencing strategy are shown in Fig. 3. Fig. 4 shows the nucleotide sequence of the noncoding strand. In the region from -250 to -120, the underlined sequence comprising of 16 nucleotides appeared 3 times, in which there were mismatches of three or less nucleotides. Such repeated sequences have also been found near the transcriptional initiation site of

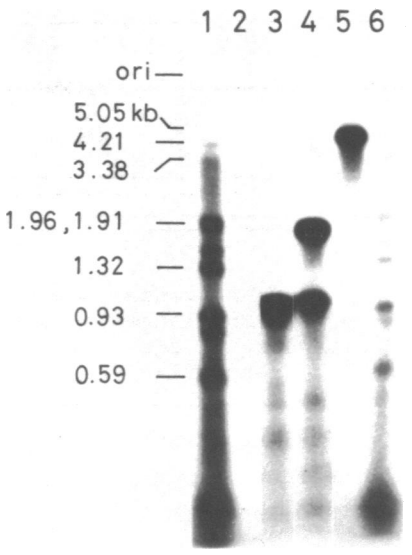

Fig. 2. Rough determination of the 5' ends of the rRNA precursor and 17S rRNA on the 4.5 kb Eco RI/Sal I fragment by S1 nuclease protection mapping. Whole nuclear RNA or purified 17S rRNA was hybridized with the coding strand of the Eco RI/Sal I fragment labeled at the 5' end with ^{32}P, digested with 1,500 units of S1 nuclease at 45°C for 30 min and electrophoresed on 1.5 % alkaline agarose gel after denaturation. Lanes 1 and 6, size markers (Eco RI/Hind III double-digested λ-DNA); lane 2, *E. coli* tRNA as a control; lane 3, 17S rRNA; lane 4, whole nuclear RNA; lane 5, end-labeled Eco RI/Sal I fragment.

rRNA genes in other eukaryotes (12,13,21), but little is known about their biological roles.

Transcriptional initiation point of the rRNA gene

To precisely map the transcription starting point on rDNA, the 260 b Ava II/Ava II fragment was prepared from the Eco RI/Sal I fragment and labeled at the 5' ends with ^{32}P, and the end-labeled coding strand was hybridized with whole nuclear RNA containing the primary transcript of the rRNA gene, followed by S1 nuclease treatment. The DNA fragment protected from S1 nuclease digestion was electrophoresed on 10 % sequencing gel in parallel with the same coding strand cleaved by nucleoside-specific chemical reactions.

Fig. 5 indicates that the 3' end of the major S1-resistant DNA fragment was located at a point 104 nucleotides upstream from

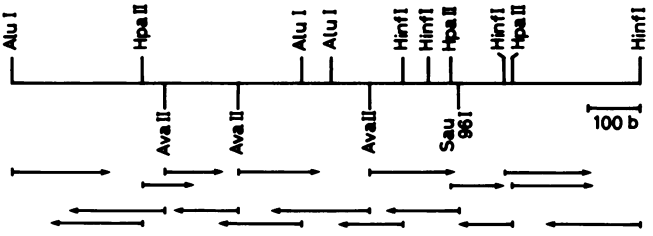

Fig. 3. Restriction map of the sequenced Alu I/Hinf I fragment and the sequencing strategy. The length of the horizontal lines and the arrows indicate the size of sequenced DNA fragments and the direction of sequencing, respectively.

the labeled 5' end of the coding strand and that the 3'-end nucleotide corresponded to the A residue on the noncoding strand when the correction of $1\frac{1}{2}$ nucleotides in length was made for the protection band according to the results of Sollner-Webb and Reeder (22). This position was therefore assigned to the transcriptional initiation point of the rRNA gene and the number +1 was gived to it (Fig. 4). However, it must be emphasized here that no test was done to determine whether the 5' end of RNA hybridized with the end-labeled DNA probe was a nucleoside triphosphate. Therefore, strictly speaking, no one can say that the RNA molecule hybridized with the DNA probe was the primary transcript of the rRNA gene. But, it is highly probable that this RNA species was the primary transcript, because no protection band larger than that shown in Fig. 5 could be found. Thus, we can tentatively conclude that transcription of the rRNA gene will start from A at position +1 on the noncoding strand in Fig. 4. This result is consistent with that of Batts-Young and Lodish who showed that the 5' end of the precursor to *D. discoideum* 17S rRNA was pppA- (23).

Comparison of the nucleotide sequences around the transcriptional initiation sites of three lower eukaryotes

Up to now, little is knowm about the consensus sequence involved in transcription of eukaryotic rRNA genes. With rRNA genes from several closely related species, it has been suggested that the sequence just around the transcriptional initiation point was highly conserved. In three species of *Xenopus*, the

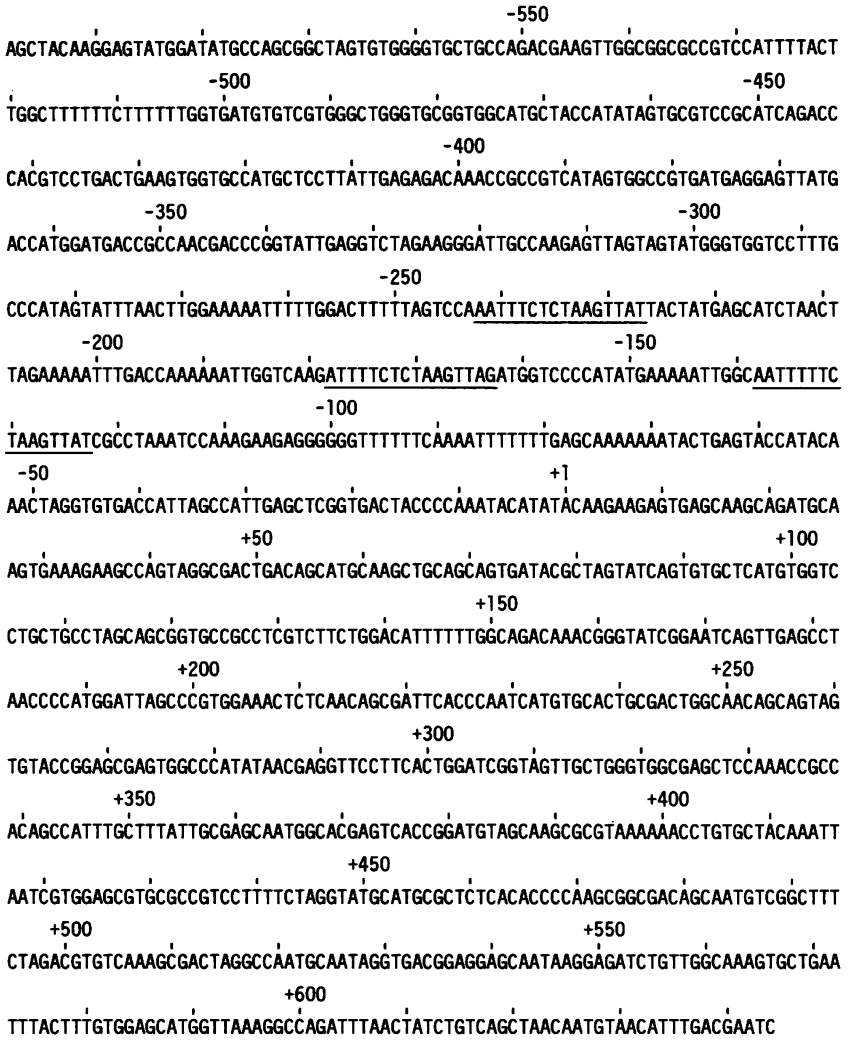

Fig. 4. Nucleotide sequence of the noncoding strand of the region surrounding the transcriptional initiation site of the *D. discoideum* rRNA gene. Position +1 is the transcriptional initiation point. Underlined sequences are the repeated sequences.

same stretch of 13 nucleotides is commonly present around the initiation point (21). Similar evidence was recently reported with three mammalian (mouse, rat and human) rRNA genes by Financsek *et al.* (24).

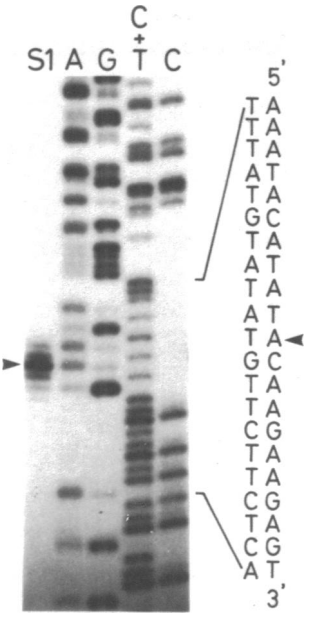

Fig. 5. Transcriptional initiation point of the *D. discoideum* rRNA gene. The 5'-end labeled coding strand of the 260 b Ava II/Ava II fragment (-159 to +101) was hybridized with whole nuclear RNA, digested with 4,500 units of S1 nuclease at 37°C for 30 min and electrophoresed on 10 % sequencing gel in parallel with the same 5'-end labeled coding strand cleaved by the nucleoside-specific chemical reactions. The arrow-head indicates the major band of the DNA fragment protected from S1 nuclease digestion.

We then compared the nucleotide sequences surrounding the transcriptional initiation site of rRNA genes in three lower eukaryotes, *D. discoideum*, *S. cerevisiae* (11) and *T. pyriformis* (12,13), with the aid of a computer. The computer analysis revealed that there is little similarity in the nucleotide

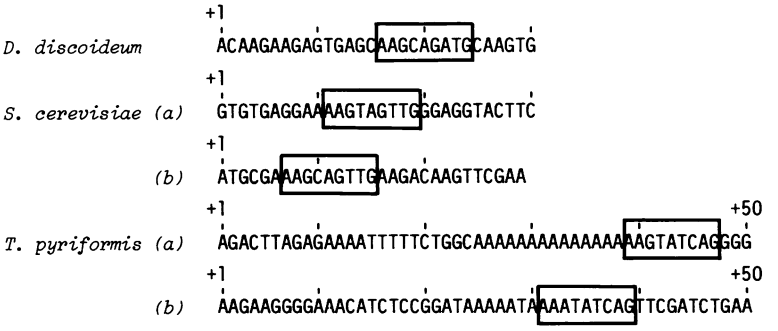

Fig. 6. Comparison of the nucleotide sequences of the region downstream from the transcriptional initiation point of the *D. discoideum*, *S. cerevisiae* (*(a)* and *(b)*, 11) and *T. pyriformis* (*(a)*, 12; *(b)*, 13) rRNA genes. The sequences showing a relatively high homology are boxed. +1 indicates the transcriptional initiation point.

sequences of the nontranscribed spacer regions, but well conserved sequences exist in the transcribed spacer regions near the initiation point. Bayev *et al.* (11) reported that in *S. cerevisiae* there were two transcriptional initiation points about 30 nucleotides apart. The sequences around these two initiation points were similar to each other. In addition, highly conserved sequences of AAG$^{T}_{C}$AGTTG were found in the region about 10 nucleotides downstream from each initiation point (Fig. 6). In *D. discoideum*, we can see a similar sequence, AAGCAGATG, in the region from +16 to +24, downstream from the initiation point (Fig. 6). The transcriptional initiation point of the *T. pyriformis* rRNA gene has been reported from two different laboratories (12,13). The major initiation point determined by Niles *et al.* (12) is located 67 nucleotides upstream from that by Saiga *et al.* (13). The nucleotide sequences around these two initiation points are not identical to each other, but highly conserved sequences of AA$^{G}_{A}$TATCAG can be seen in the region about 40 nucleotides downstream from each initiation point (Fig. 6). These sequences are more similar to those in *S. cerevisiae* than that in *D. discoideum*.

The above findings seem to suggest that the conserved sequences may act as one of the signals specifying the initiation of transcription. But, there is no direct evidence of the positive roles of the conserved sequences in transcription of rRNA genes as yet.

ACKNOWLEDGEMENTS

We thank Dr. R. A. Firtel for providing the recombinant plasmid, pDd 507. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Grummt,I. (1981) Proc. Natl. Acad. Sci. USA 78, 727-731.
2. Grummt,I. (1981) Nucleic Acids Res. 9, 6093-6102.
3. Mishima,Y., Yamamoto,O., Kominami,R. and Muramatsu,M. (1981) Nucleic Acids Res. 9, 6773-6785.
4. Miller,K.G. and Sollner-Webb,B. (1981) Cell 27, 165-174.
5. Kohorn,B.D. and Rae,P.M.M. (1982) Proc. Natl. Acad. Sci. USA 79, 1501-1505.

6. Grummt,I., Roth,E. and Paule,M.R. (1982) Nature 296, 173-174.
7. Cockburn,A.F., Taylor,W.C. and Firtel,R.A. (1978) Chromosoma 70,19-29.
8. Iwabuchi,M., Mizukami,Y. and Sameshima,M. (1971) Biochim. Biophys. Acta 228, 693-700.
9. Batts-Young,B., Maizels,N. and Lodish,H.F. (1977) J. Biol. Chem. 252, 3952-3960.
10. Cockburn,A.F., Newkirk,M.J. and Firtel,R.A. (1976) Cell 9, 605-613.
11. Bayev,A.A., Georgiev,O.I., Hadjiolov,A.A., Kermekchiev,M.B., Nikolaev,N., Skryabin,K.G. and Zakharyev,V.M. (1980) Nucleic Acids Res. 8, 4919-4926.
12. Niles,E.G., Sutiphong,J. and Haque,S. (1981) J. Biol. Chem. 256, 12849-12856.
13. Saiga,H., Mizumoto,K., Matsui,T. and Higashinakagawa,T. (1982) Nucleic Acids Res. 10, 4223-4236.
14. Clewell,D.B. and Helinski,D.R. (1969) Proc. Natl. Acad. Sci. USA 62, 1159-1166.
15. Maxam,A.M. and Gilbert,W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
16. Smith,H.O. and Birnstiel,M.L. (1976) Nucleic Acids Res. 3, 2387-2399.
17. Berk,A.J. and Sharp,P.A. (1977) Cell 12, 721-732.
18. McDonell,M.W., Simon,M.N. and Studier,F.W. (1977) J. Mol. Biol. 110, 119-146.
19. Yagura,T., Yanagisawa,M. and Iwabuchi,M. (1976) Biochem. Biophys. Res. Commun. 68, 183-189.
20. Frankel,G., Cockburn,A.F., Kindle,K.L. and Firtel,R.A. (1977) J. Mol. Biol. 109, 539-558.
21. Bach,R., Allet,B. and Crippa,M. (1981) Nucleic Acids Res. 9, 5311-5330.
22. Sollner-Webb,B. and Reeder,R.H. (1979) Cell 18, 485-499.
23. Batts-Young,B. and Lodish,H.F. (1978) Proc. Natl. Acad. Sci. USA 75, 740-744.
24. Financsek,I., Mizumoto,K., Mishima,Y. and Muramatsu,M. (1982) Proc. Natl. Acad. Sci. USA 79, 3092-3096.