Volume 8 Number 8 1980 Nucleic Acids Research

Nucleotide sequence of a 5S ribosomal RNA gene in the sea urchin Lytechinus variegatus

A-Lien Lu[†], Deborah A.Steege* and Darrel W.Stafford[+]

Departments of Zoology[+] and Biochemistry[†], University of North Carolina, Chapel Hill, NC 27514, and *Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

Received 2 January 1980

ABSTRACT

The nucleotide sequence of a cloned DNA fragment corresponding to a gene for 5S ribosomal RNA from the sea urchin Lytechinus variegatus has been determined. This sequence is representative of the dominant species of 5S rRNA labelled in vivo with $^{32}PO_4$ during the cleavage stage of Lytechinus embryonic development.

INTRODUCTION

The nucleotide sequences of 5S ribosomal RNAs have been determined for a number of prokaryotic and eukaryotic organisms, and appear to be highly conserved throughout evolution (1-3). In all eukaryotes studied thus far, the genomic DNA encoding 5S rRNA (4) is found in multigene families in tandem arrays (5-11). In Xenopus, for example, the rRNA genes occur in two classes, one of which is expressed in oocyte cells, and the other, in somatic tissue (3,12). Current work on the structural organization and expression of 5S rRNA genes has been stimulated by the development of in vitro transcription systems which allow accurate transcription of single 5S rDNA repeats (13-16). The fidelity of this transcription suggests that each 5S rRNA gene includes the information for initiation and termination of transcription, and raises the possibility that additional regulatory signals may also be present.

Neither primary sequence data nor evidence for the genomic organization of 5S rDNA has been obtained for the sea urchin, an echinoderm which has been used extensively for studies of differential gene expression during embryonic development. Nijhawan and Marzluff (17) have recently reported that in the sea urchin Lytechinus variegatus there are about 175 copies per haploid genome of the gene encoding 5S rRNA. We have found that the Lytechinus 5S rDNA repeat contains a single recognition site for the restriction endonuclease Hae III. Cleavage of 5S rDNA with this enzyme yields fragments with chain lengths of 1.2 and 1.3 kilobases (18). The pattern of partial diges-

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

tion products we obtain from Hae III cleavage of total L. variegatus DNA is consistent with an arrangement of 5S rRNA genes in at least two tandemly repeated, non-interspersed families. We have cloned Eco RI fragments of Lytechinus DNA enriched in 5S rDNA sequences into the plasmid vector pACYC 184 (18). As a first step in our studies of sea urchin 5S rRNA gene expression, we present in this report the nucleotide sequence for the region of the Lytechinus DNA insert corresponding to a 5S rRNA gene. We show also that the 5S rRNA synthesized during the cleavage stage of development yields RNase T1 and RNase A oligonucleotides which are consistent with the cloned 5S rDNA sequence.

MATERIALS AND METHODS

Materials

Dimethyl sulfate and diethylpyrocarbonate, hydrazine, and piperidine were from Aldrich, Eastman Kodak, and Fisher, respectively. [γ-^{32}P]ATP (3000 Ci/mmol) was purchased from New England Nuclear, and carrier-free [^{32}P]orthophosphoric acid, from Amersham. The following chromatographic materials were used for RNA sequencing: polyethyleneimine (PEI)-cellulose thin layer plates (Brinkmann), DEAE paper (DE 81, Whatman), Avicel thin layer plates (Analtech), and Cellogel strips (Kalex). Calf thymus alkaline phosphatase was a gift from I. Sures, Stanford University. The restriction endonuclease Hae III was purified by the general procedure of Bickel et al. (19). Eco RI was purified from E coli strain RY13(R1) as described by Modrich and Zabel (20). Other enzymes were obtained commercially: T4-polynucleotide kinase (P-L Biochemicals); Alu I, Hpa II, Mbo II, and Hha I (New England Biolabs); RNases T1, T2, and U_2 (Calbiochem); RNase A (Worthington); and snake venom phosphodiesterase (Boehringer-Mannheim).

Preparation of Sea Urchin 5S DNA fragments

Sea urchin (Lytechinus variegatus) DNA, prepared as described previously (21), was enriched in 5S DNA sequences by several steps of actinomycin D-cesium chloride equilibrium density gradient centrifugation (7). DNA preparations were digested with the restriction endonuclease Eco RI, and the resulting fragments were ligated to pACYC 184 DNA cleaved with Eco RI (22). Clones containing plasmids with 5S rDNA inserts were identified among the tetracycline-resistant, chloramphenicol-sensitive transformants (23) by hybridization with purified [^{125}I]-sea urchin 5S rRNA (24). Of four clones carrying plasmids with 5S rDNA-containing insertions of about the same size, one denoted pLu

103 has been used for further studies. The position of the 5S rDNA within the 1.34 kb insert in pLu 103 was determined by hybridizing ^{125}I -labelled sea urchin 5S rRNA to the restriction fragments generated by the endonucleases Alu I, Eco RI, Hae III, Hinf I, Hha I, and Hpa II. Details of the cloning experiments and restriction endonuclease mapping will be presented elsewhere (18). The cloning of these recombinant DNA molecules and all subsequent propagation of the pLu 103 plasmid in E coli strain JC1569 were conducted according to the NIH Guidelines (P2/EK1).

Plasmid DNA from pLu 103 was prepared as described previously (25). Following cleavage with Hae III, the Hae III fragment of approximately 1.02 kb (Fig. 1) was purified by polyacrylamide gel electrophoresis, and redigested with the following restriction enzymes to obtain fragments suitable for DNA sequencing: Hpa II, Alu I, Mbo II and Hha I. These restriction enzymes were used according to the suppliers' recommendations.

DNA fragments were fractionated by electrophoresis on 3 cm tubes or 19x19x0.3 cm slab gels of 6% acrylamide (29:1 acrylamide:bisacrylamide)/50 mM Tris-borate, pH 8.3/1 mM EDTA. The positions of DNA fragments in the gel were identified by UV fluorescence after staining with ethidium bromide (1 µg/ml) or by autoradiography. Gel slices were subjected to electrophoretic elution (26) at 180V for a minimum of three hours in 5 mM Tris-HCl, pH 7.4/ 2.5 mM sodium acetate/0.1 mM EDTA. Residual gel particles were removed from the eluate by its filtration through a layer of glass wool in a syringe. Magnesium acetate and ammonium acetate were added to DNA solutions to give final concentrations of 0.01 M and 0.1 M, respectively, and the DNA was precipitated with 2 volumes of ethanol. After 15 minutes in a dry-ice-ethanol bath and centrifugation at 12,000 x g for 15 minutes, the precipitated DNA was dissolved in 0.3 M sodium acetate, pH 6.5, and reprecipitated as above.

5'-Terminal Labelling of DNA Fragments

DNA samples, dissolved in 90 µl H_2O, were brought to 0.1 M Tris-HCl, pH 8.0, and treated with 0.02 units of calf thymus alkaline phosphatase for one hour at 37°C. After adding 0.3 ml of 0.1 M Tris-HCl, pH 8.0, the reaction mixture was extracted twice with phenol and then once with ether. Following ethanol precipitation as above, end labelling was carried out as described by Maxam and Gilbert (27), using 10 units of polynucleotide kinase for 45 minutes at 37°C per 30-50 pmole sample of single-stranded DNA.

Strand Separation

The strands of the 76 bp Hpa II/Hpa II fragment and the 249 bp Mbo II/

Mbo II fragment (Fig. 1) were separated using minor modifications of the procedure described by Maxam and Gilbert (27). The double-stranded fragments, labelled at their 5' ends with ^{32}P, were denatured by heating for three minutes at 90°C in 40 µl of 10 mM Tris-borate pH 8.3/2 mM EDTA/0.002% xylene cyanole FF/0.002% bromphenol blue/80% dimethyl sulfoxide/2.4% glycerol. The sample was chilled quickly to 0°C, and loaded immediately in 3.2 cm slots in a slab gel (18x40x0.25 cm) of 5% acrylamide (60:1 acrylamide: bisacrylamide)/50 mM Tris-borate, pH 8.3/1 mM EDTA. Electrophoresis at 130V was carried out at 4°C for 14 hours to prevent renaturation of the separated strands, and then continued at room temperature once the DNA molecules had moved well into the gel. Electrophoretic elution of the separated strands was as described above. After ethanol precipitation of the single stranded DNAs, using 35 µg of unlabelled sonicated calf thymus DNA as carrier, final resuspension was in 20 µl of water.

DNA Sequencing

Nested sets of 5' end-labelled fragments were produced and fractionated by electrophoresis on polyacrylamide gels as described in the chemical degradation DNA sequencing method of Maxam and Gilbert (27), with minor modifications. An additional ethanol precipitation was included in the protocol prior to piperidine cleavage to ensure complete removal of the hydrazine. Aliquots of the products generated by the four reactions G, A>G, C and C+T were fractionated on 14% acrylamide gels containing 7 M urea. After electrophoresis at constant voltage (700-900 V), gels were covered with Saran Wrap and exposed to flash-sensitized Kodak XR-1 X-ray film at -20°C (28).

Isolation and Analysis of [^{32}P]5S rRNA Labelled in Vivo

At the blastula stage of development, 20 mCi of $^{32}PO_4$ was added to a 1-liter solution of phosphate-free sea water (29) containing about 5.5×10^6 Lytechinus variegatus embryos which had been fertilized 5 hours earlier. Development was allowed to continue, and at the pluteus stage, the embryos were collected on 44 micron Nitex screens (Tetko), washed with cold water, and transferred to centrifuge tubes. After centrifugation for 10 min at 12,000 x g, the pelleted embryos were washed once with 25 mM Tris-HCl, pH 7.5/ 25 mM NaCl/1 mM spermidine/1 mM dithiothreitol/0.5 mM phenylmethyl sulfonyl fluoride/2% Triton X-100, and resuspended in 7 volumes of the same buffer. Following 15 strokes of a tight pestle in a 40 ml dounce homogenizer, the homogenate was centrifuged at 12,000xg for 10 minutes. Ribosomes were isolated from the 27,000 x g supernatant as described by Palmiter (30). Ribosomes

were dissociated in 0.05 M sodium acetate, pH 5.0/0.1 M NaCl/ 1% sodium dodecyl sulfate, and proteins removed by four extractions with equal volumes of phenol. After ethanol precipitation, the RNA was dissolved in 10 ml of water. NaCl was added at a final concentration of 4 M, the sample placed at -20°C overnight, and the high molecular weight RNA removed by centrifugation for 15 minutes at 17,000 x g. This supernatant was dialyzed against 1 liter of water which had been treated with diethylpyrocarbonate prior to autoclaving, and reprecipitated with 2 volumes of ethanol. 5S rRNA was purified by electrophoresis on slabs of 10% acrylamide (29:1 acrylamide: bisacrylamide)/50 mM Tris-borate, pH 8.3/1 mM EDTA, and eluted electrophoretically as described above. The uniformly-labelled [^{32}P]5S rRNA was analyzed using standard Sanger sequencing techniques (31). Secondary RNase U_2 digestion products were fractionated by electrophoresis in pryidinium acetate, pH 3.5, on Whatman DE81 paper, and mononucleotides were in general separated by chromatography on thin layers of Avicel (32).

RESULTS AND DISCUSSION

A Lytechinus gene for 5S rRNA

From our initial hybridizations of ^{125}I-labelled sea urchin 5S rRNA to restriction fragments derived from the recombinant plasmid pLu 103, we obtained evidence that a nucleotide sequence corresponding to 5S rRNA was contained in an *Alu* I fragment with a chain length of about 380 base pairs (18). Based on the position of this fragment in the pLu 103 restriction map, we selected for sequence analysis several sets of restriction fragments including this region of the sea urchin DNA insert which would provide complementary data from both DNA strands and give appropriate overlaps between restriction fragments. These are shown schematically in Figure 1. Nucleotide sequence analysis of these fragments was carried out using the chemical degradation methods of Maxam and Gilbert (27). From the data obtained, we are able to deduce the sequence of 248 base pairs shown in Fig. 2. Centered within this sequence is a region (designated 1-120) which has a high degree of homology with other eukaryotic 5S rDNAs. In particular, this sequence differs by only 16 nucleotides from *Drosophila* 5S rDNA and by 18 nucleotides from the *Xenopus laevis* oocyte 5S rDNA.

There is currently no 5S rRNA sequence information for the echinoderm phylum of which the sea urchin *Lytechinus variegatus* is a member. To determine whether the 5S rRNA-like sequence in pLu 103 is represented in *in vivo* 5S rRNA, we labelled developing embryos with $^{32}PO_4$ and analyzed the purified 5S

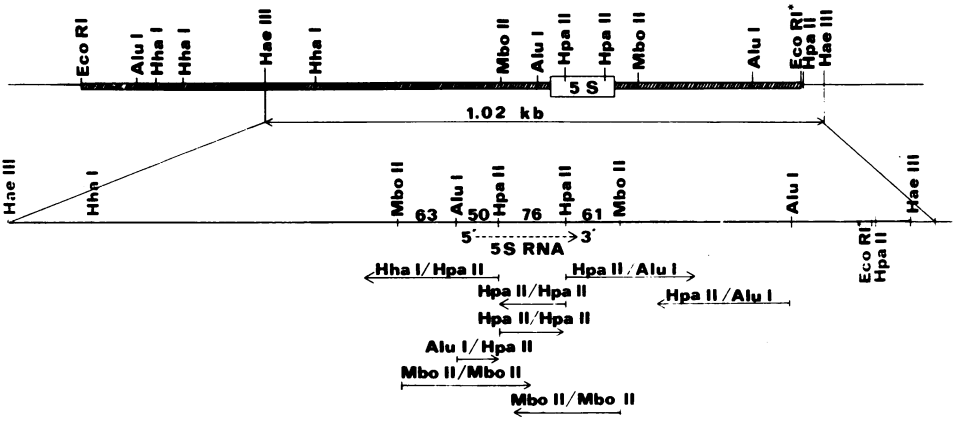

Figure 1: Restriction map and sequence strategy for the analysis of the Lytechinus variegatus insert (▨) in the pLu 103 plasmid. The coding region for 5S rRNA is indicated by the open box. In the more detailed diagram for the map between Hae III restriction sites, which includes the 5S rDNA sequence (dotted line), numbers refer to restriction fragment chain lengths, in base pairs, determined from the DNA sequence. The 5' to 3' polarity and regions of DNA single strands analyzed by the DNA sequencing method of Maxam and Gilbert (27) are indicated by arrows. Given above each arrow is the restriction fragment, designated by its left and right cleavage ends, from which the sequence information was obtained. The set of overlapping restriction fragments described below were generated from a Hae III fragment of about 1.02 kb which we had found to contain sequences hybridizable to [^{125}I]5S rRNA. The Hae III fragment was exised from the plasmid as described in the Materials and Methods. The three largest fragments obtained from this after Hpa II digestion were purified by acrylamide gel electrophoresis and labelled with ^{32}P at their 5' ends. The singly labelled Hha I/Hpa II and Hpa II/Alu I fragments were then generated by cleavage of Hae III/Hpa II and Hpa II/Hpa II (∿350 bp) with Hha I and Alu I, respectively. The strands of the 76 bp Hpa II/Hpa II fragment were separated as described in the Materials and Methods, and each strand was subjected to chemical sequencing. Following cleavage of the 1.02 kb Hae III fragment with Alu I and subsequent polynucleotide kinase treatment of a 380 bp fragment, the fragments Alu I/Hpa II and Hpa II/Alu I, labelled uniquely at their Alu I ends, were generated by Hpa II cleavage. Each strand of a 5'-end labelled 249 bp Mbo II product obtained from the 1.02 kb Hae III fragment was analyzed.

rRNA by standard Sanger RNA sequencing methods (31). The two-dimensional RNase T1 and RNase A fingerprints we obtained are shown in Fig. 3. The compositions of all oligonucleotides determined from further analysis (Tables I and II) are in full agreement with those predicted by the pLu 103 DNA sequence. These results suggest that the cloned rDNA sequence is representative of the dominant species of 5S rRNA synthesized in vivo at this stage

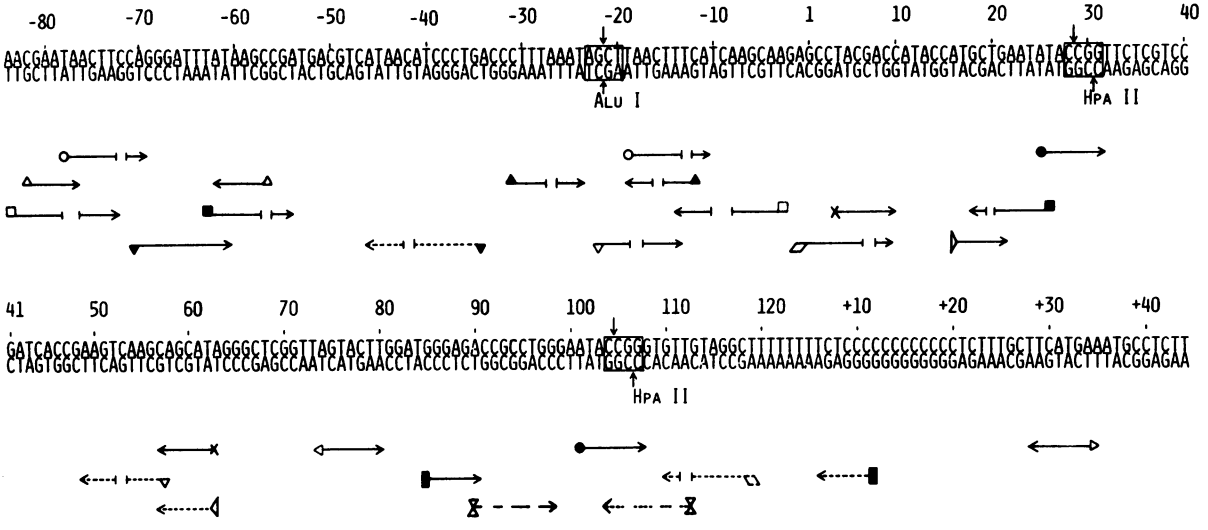

Figure 2: Sequence of a Lytechinus variegatus 5S rRNA gene. Nucleotides preceding the region corresponding to the mature 5S rRNA species are denoted -84 to -1; the coding region is numbered 1-120; and the residues beyond the CTT sequence corresponding to the 5S rRNA 3' terminus are designated +1 through +44. The recognition sequences for restriction endonucleases are enclosed in boxes, with cleavage sites indicated by vertical arrows. The region is bounded by Mbo II sites which occur 8 nucleotides beyond the right and left ends shown. Direct repeats (⟶ ⟶), symmetric regions (⟶ ⟵) and dyad symmetries (⟶ ⟵---) are indicated in the appropriate positions below the sequence. Imperfect matches are represented by interruptions (⟶⊣ ⊢⟵) in the arrow stems.

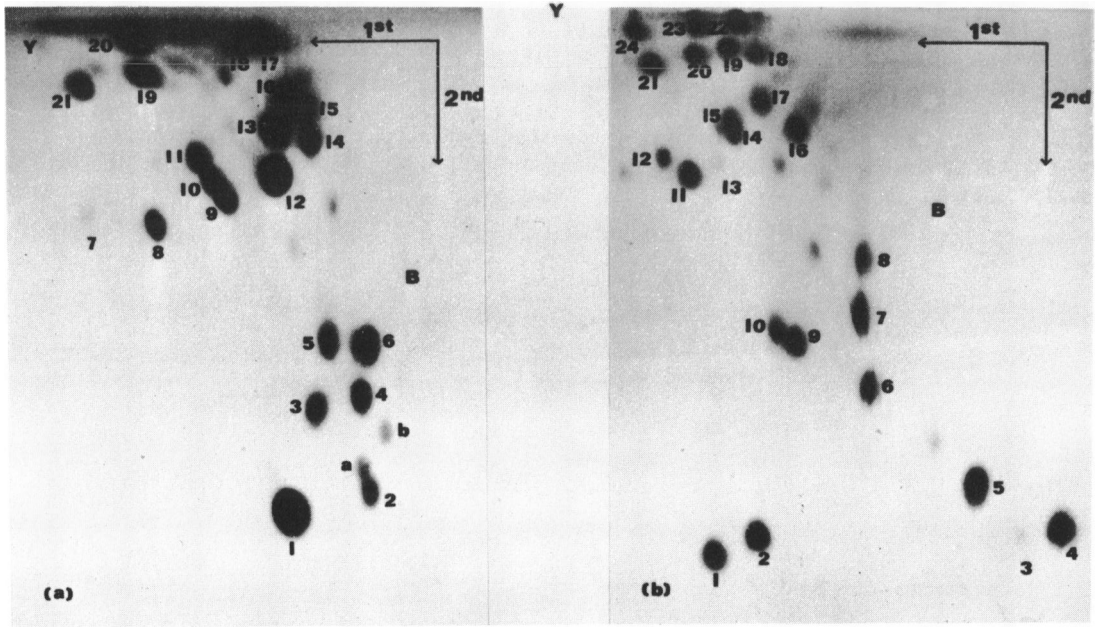

Figure 3: Two-dimensional fingerprints of a) T1 RNase-digested and b) RNase A-digested ^{32}P-labelled sea urchin embryo 5S rRNA. Oligonucleotide numbers correspond to the data presented in Tables I and II. ^{32}P-labelled 5S rRNA was purified as described in the Materials and Methods. Electrophoresis was on cellulose acetate (Cellogel) in 7 M urea/pyridinium acetate, pH 3.5, for the first dimension and in 7% formic acid on DEAE paper for the second. B and Y mark the positions of the blue (xylene cyanole FF) and yellow (orange G) dye markers. Arrows indicate the directions of the first and second dimensions; the origin, however, is not included in the sections of the autoradiograms photographed. As part of the search for oligonucleotides derived from the 5' and 3' termini of 5S rRNA, all minor oligonucleotides visible in the fingerprints, but not numbered, were analyzed. With the exception of T2, T7, and P13, which are discussed in the text, these minor oligonucleotides did not have the properties appropriate for 5' and 3' ends.

TABLE I

Sea Urchin 5S rRNA T1 RNase Oligonucleotides

Oligonucleotide Number	RNase A Digestion Products	RNase U_2 Digestion Products	Oligonucleotide Sequence Predicted from DNA Sequence	Experimental Molar Yield	Expected Yield
T1	Gp		Gp	8.3	12
T2[a]	Cp, Up(1.0:1.0)		C-U-U_{OH}	0.8	1
T3	A-Gp		A-Gp	1.4	1
T4	Cp(0.9), A-Gp(1)		C-A-Gp	0.9	1
T5	A-A-Gp		A-A-Gp	0.8	1
T6	A-Cp(0.8), Cp(1.3), Gp(1)	Ap(0.9), C-C-Gp(1)	A-C-C-Gp	1.0	1
T7[b]	pGp		pGp	0.2	-
T8	Up, Gp(~1:1)		U-Gp	1.2	1
T9	Cp, Up, Gp(~1:1:1)		C-U-Gp	0.9	1
T10	Up, A-Gp (~1:1)		U-A-Gp	0.8	1
T11	A-Up,Gp (~1:1)		A-U-Gp	1.0	1
T12	Cp(1.6), Up(0.9), Gp(1)		C-U-C-Gp, U-C-C-Gp, C-C-U-Gp	2.6	3
T13	Cp(1.9), Up(0.7), A-Up(1.1), A-A-Gp (1.1), A-Gp(1)	C-Ap(1.1),U-Ap(1.3) (U,C)Ap(0.4) (U,C)A-Ap(0.5) Ap+Gp(3)	U-C-A-A-Gp C-A-U-A-Gp	1.6	2
T14	Cp(1.5), Up(1.0), A-Cp(1.1), Gp(1)	(C_2,U)Ap(1.8), C-Gp(1)	C-C-U-A-C-Gp	0.6	1
T15	Cp(1.2),A-Up(1.1), A-Cp(1.1),Gp(1)	Ap(0.9),(U,C)Ap(1.2) C-C-Gp(1)	A-U-C-A-C-C-Gp	0.7	1
T16	Cp(0.8),A-Cp(1.0) A-A-Up(1.0),G(1)	Ap,U-Ap,A-Ap,C-C-Gp	A-A-U-A-C-C-Gp	0.8	1
T17	Cp(1.4),A-Cp(1.8), A-Up(2.1),Gp(1)	Ap,U-Ap,C-C-Ap, U-Gp	A-C-C-A-U-A-C-C-A-U-Gp	0.7	1
T18	Cp(1.1),A-Cp(0.8), A-Up(1.0),A-A-Up(1.1) Gp(1)	Ap(0.9),A-Ap(0.8) U-Ap(2.5),C-C-Gp(1)	A-A-U-A-U-A-C-C-Gp	0.7	1
T19	U(1.8),A-Gp(1)	U-U-Ap,Gp	U-U-A-Gp	1.2	1
T20[c]	Cp(1.3),Up(5.0), A-Cp(1.1),Gp(2)	U-Ap,(U_2,C_1)Gp, (U_x,C_y)Gp	U-U-C-U-C-Up U-A-C-U-U-Gp	1.7	2
T21	Up(1.8),Gp(1)		U-U-Gp	1.0	1
Ta[d]	Cp(0.8),Gp(1)		C-Gp	0.6	0
Tb[d]	Cp(1.4),Gp(1)		C-C-Gp	0.3	0

Table I. RNase T1 oligonucleotides from [^{32}P]-labelled sea urchin embryo 5S rRNA were analyzed using Sanger sequencing methods (31). Oligonucleotide numbers correspond to Fig. 3a. The relative molar yields of primary oligonucleotides were normalized to the yield of ACCG(T6). Expected yields are based on the region of the pLu 103 DNA sequence proposed to encode a 5S rRNA. Ratios among secondary RNase A and RNase U_2 digestion products are expressed relative to the yields of the Gp-containing products. In some cases approximate ratios were determined (~) by visual inspection. Where an underdigestion product obtained with U_2 RNase helps to order a T1 RNase oligonucleotide sequence, it has been included in the Table. The compositions of each product were confirmed by complete digestion with RNase T_2 and subsequent fractionation of mononucleotides by thin layer chromatography (32).

[a]The single product obtained from complete digestion of oligonucleotide T2 with snake venom phosphodiesterase was pU.

[b]T7 was identified as pGp by its mobility on DE81 paper in pyridinium acetate, pH 3.5 and on thin layer Avicel plates in the isobutyric acid-NH_3 and t-butanol-HCl solvent systems (32).

[c]Digestion of T20 with U_2 RNase gives U-Ap and two oligonucleotides which have mobilities, relative to markers, that are consistant with lengths 4 nucleotides and 6 nucleotides, respectively.

of development.

The oligonucleotide maps of the ^{32}P-labelled in vivo 5S rRNA also provide information about the 5' and 3' ends of the mature 5S rRNA. The properties of three oligonucleotides are those expected for the 5' and 3' termini. The RNase

TABLE II

Sea Urchin 5S rRNA RNase A Oligonucleotides

Oligonucleotide Number	RNase T1 Digestion Products	Oligonucleotide Sequence Predicted from DNA Sequence	Experimental Molar Yield	Expected[3] Yield
P1, P2		U>p,Up	15.4	11
P3,P4		C>p,Cp	13.2	15
P5		A-Cp	5.0	6
P6	Gp(1.2),Cp(1)	G-Cp	3.0	2
P7	A-Gp(0.9), A-Cp(0.7),Gp(1.0), Cp(1)	G-A-Cp, A-G-Cp	2.0	2
P8	A-A-Gp(1.4), Cp(1)	A-A-G-Cp	1.0	1
P9,P10	A-Up	A-Up, A-U>p	5.6	4
P11,P12	Gp(1.2), Up(1)	G-Up, G-U>p	4.8	3
P13[a]		pG-Cp	0.1	–
P14	Gp(0.9), A-Up(1)	G-A-Up	1.4	1
P15	A-Gp(1.1), Up(1)	A-G-Up	1.4	1
P16	A-Gp(1.6), Gp(1.5), Cp(1)	A-G-G-Cp	1.4	1
P17	Gp(0.8), A-A-Up(1)	G-A-A-Up	1.2	1
P18	A-Gp(1.1), Gp(1.8), Cp(1)	A-G-G-G-Cp	0.9	1
P19	A-A-Gp(1.4), Gp(1.5), Up(1)	G-A-A-G-Up	1.0	1
P20	Gp(1.4), A-Up(1)	G-G-A-Up	1.5	1
P21	Gp(2.8), Up(1)	G-G-Up	3.0	2
P22	A-Gp(1.3), Gp(3.6), A-Cp(1)	G-G-G-A-G-A-Cp	0.8	1
P23	Gp(2.2), A-A-Up(1)	G-G-G-A-A-Up	1.0	1
P24	Gp(3.1), Up(1)	G-G-G-Up	1.1	1

Table II. RNase A oligonucleotides obtained from sea urchin 5S rRNA labelled with $^{32}PO_4$ in vivo were characterized using standard Sanger sequencing methods (31). Oligonucleotide numbers correspond to the fingerprint of Fig. 3b. The molar yields of primary RNase A oligonucleotides were normalized to the yield of AAGC (P8), which occurs uniquely in the sequence. Expected yields are based on the pLu 103 5S rDNA sequence. Yields of secondary RNase T1 digestion products are expressed relative to the pyrimidine-containing product. Oligonucleotide compositions were confirmed by digestion to the mononucleotide level with RNase T2, followed by thin layer chromatography (32).

[a]The T2 digestion products obtained from P13, as identified by thin layer chromatography in the isobutyric acid-NH_3 and t-butanol-HCl solvent systems (31), were pGp and Cp.

T1 product designated T2 does not contain the G residue characteristic of all RNase T1 digestion products derived from internal regions of an RNA molecule. When digested to completion with RNase T2, which generates nucleoside 3' monophosphates, oligonucleotide T2 gives Cp and Up in a 1:1 ratio. Parallel digestion with snake venom phosphodiesterase, which produces nucleoside 5' monophosphates, gives only pU. These data, taken together with the observation that a nucleoside diphosphate (pXp) is not found among the snake venom phosphodiesterase products, indicate that oligonucleotide T2 has the sequence C-U-U_{OH} and represents the 5S rRNA 3' terminus. A corresponding CTT sequence is present at position 118-120 in the pLu 103 DNA sequence (Fig. 2). Our oligonucleotide data (Fig. 3a, Table I) and, in addition, preliminary analysis by chemical sequencing methods of the sea urchin 5S rRNA labelled in vitro with

[^{32}P]pCp (33) (A. Lu and D. Stafford, unpublished results) give no evidence for length heterogeneity at the 3' end. The other two oligonucleotides of interest, T1 RNase product T7 and RNase A product P13, both contain ribosyl 5'-monophosphate moieties derived from the 5' terminus. Within the limits of our analysis, these are the only ribosyl 5' monophosphates detectable in RNase T1 and RNase A fingerprints. From appropriate analysis, T7 is pGp and P13, pGpCp. Since the pLu 103 DNA sequence contains the GC dinucleotide at two positions, at -6 and 1 (Fig. 2), the pancreatic RNase oligonucleotide pGpCp does not unambiguously define the 5' terminal residue. From chemical sequencing analysis of 5S rRNA labelled *in vitro* with [^{32}P]pCp (33), however, we are able to conclude that the 5S rRNA does not extend to the -6 position (A. Lu and D. Stafford, unpublished results); hence the second GC (1) corresponds to the RNA 5' terminus.

The 5' monophosphorylated oligonucleotides T7 and P13 are observed reproducibly, although in low molar yields, in the fingerprints for several preparations of *in vivo* ^{32}P-labelled 5S rRNA. If the sea urchin 5S rRNA is a primary transcript, as appears to be the case for other eukaryotic organisms (1,13,15,16), it should have a 5' end generated directly by transcription initiation rather than by secondary cleavages during processing events. One would expect it therefore to carry a triphosphorylated nucleoside. However, monophosphorylated nucleosides have generally been found by others as well at the 5' ends of the 5S rRNAs isolated from several eukaryotic organisms (1) for which other criteria suggest that the 5S rRNAs are primary transcripts (13,15,16). We have found that our preparations of 5S rRNA from sea urchin eggs can be capped with an efficiency of up to 16% with the vaccinia virus capping enzyme (gift of B. Moss) in the presence of [α-^{32}P]GTP (N. Blin and D. Stafford, unpublished results). Since the capping enzyme preparation requires a di- or triphosphate for activity (34,35), the capping result suggests that these 5' termini are most likely represented in the population of egg 5S rRNA molecules, even though they are not detectable in fingerprints. The 5' terminal pGpCp we observe thus may be a degradation product of a nucleoside triphosphate moiety.

Features of the 5S rDNA sequence

We have examined the 248-base pair *Lytechinus* DNA sequence including the gene for 5S rRNA, and have used the computer program developed by Korn et al. (36) to search for repeated sequences, symmetric regions, and dyad symmetries that could prove significant in defining the secondary structure of the 5S rRNA and in protein interactions with 5S rDNA. The more striking examples

from each class are designated by arrows below the sequence shown in Fig. 2. We have also compared the sea urchin 5S rDNA to other eukaryotic 5S rDNAs, and find that the high degree of homology the sea urchin 5S rRNA shares with the 5S rRNA of other species does not extend to the flanking sequences. Features of the 5S rRNA gene are discussed below.

The 5' noncoding region is relatively AT-rich; 22 of the 30 base pairs preceding the RNA start site are A-T base pairs. The coding region exhibits a number of dyad symmetries which presumably reflect the existence of secondary structure in the 5S rRNA molecule. The positions of potential base-paired regions in the sea urchin sequence are consistent with secondary structures which have been proposed for eukaryotic 5S rRNAs (2).

The region of the sea urchin 5S rDNA sequence in the vicinity of the RNA 3' terminus has several features similar to sequences at the ends of other eukaryotic genes. As shown in Fig. 2, there is a stretch of 6T's followed by 12C's beyond the CTT of the sense strand. In other eukaryotic genes analyzed there are clusters of at least 4T's beyond the residue corresponding to the RNA 3' end: 4T's in the adenovirus VA RNA_I gene (37), 5T's in a Drosophila 5S rRNA gene (38), a stretch of 21 or 29 T's in yeast 5S rRNA genes (39,40), and a T-rich sequence TTTTCAACTTTATTTT in a Xenopus laevis oocyte 5S rRNA gene (41). In the sea urchin gene the AT-rich region generated by the cluster of T residues is preceded by a relatively GC-rich region; this includes a dyad symmetry (Fig. 2) which would correspond in 5S rRNA to a 7 base-pair stem closed by a loop of 5 residues. These features, a GC-rich region and dyad symmetry followed by an AT-rich region, have been observed at the ends of other eukaryotic genes (42) and at prokaryotic transcription termination sites (43). As has been observed for the other eukaryotic genes, the dyad symmetry is separated by a gap of several base pairs from the cluster of T residues. At prokaryotic termination sites, by contrast, the string of T residues is immediately adjacent to a dyad symmetry. In the sea urchin sequence a stretch of CG base pairs with a perfect purine/pyrimidine strand bias follows the T cluster. This could possibly function to prevent readthrough by RNA polymerase. A final interesting feature of this region is a purine/pyrimidine strand bias that begins at position 118 (Fig. 2) and continues as far as our sequence analysis has been extended. This appears to be distinctive of the 3' noncoding region. The autoradiograms of DNA sequencing gels shown in Figure 4 are given as illustrations of this strikingly asymmetric distribution. The single-stranded DNA sequences shown represent segments of the 3' noncoding region beyond that presented in Fig. 2. The strand bias does not occur in

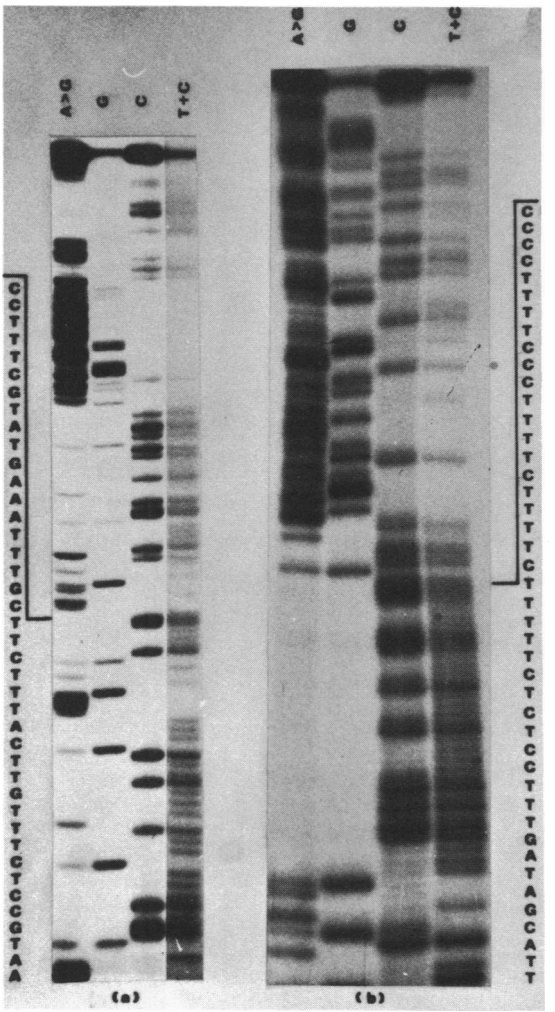

Figure 4: DNA sequence gels showing the nonrandom distribution of purines and pyrimidines discussed in the text which occurs on DNA strands in the region following the 5S rRNA coding sequence. The display shown in a, obtained from the strand of the Hpa II/Alu I fragment labelled at the 5' Hpa II end, gives the products for the region of the sense strand corresponding to approximately +35 to +75 (bottom to top). The pattern in b, obtained from the complementary strand of Hpa II/Alu I labelled at the 5' Alu I end, shows the region corresponding to approximately (bottom to top) +172 to +129.

the 140 nucleotides we have examined in the 5' noncoding region.

In the fingerprints of ^{32}P-labelled *in vivo* 5S rRNA, a few oligonucleotides which are not predicted by the DNA sequence appear in fractional molar yields. If they are not derived from contaminating RNAs, these products possibly reflect heterogeneity within the 5S rRNA species. To look for sequence heterogeneity among the sea urchin 5S rDNA coding regions and to determine whether the genomic rDNA exists in distinct arrays of genes expressed differentially in development, we are currently examining additional copies of cloned 5S rDNA repeats, and analyzing the 5S rRNAs synthesized during different developmental stages.

ACKNOWLEDGEMENTS

This work was supported by NSF grant PCM 77/03962 and NIH Grant GM 24233 to D.W.S. and by Grant # AI 14882 from the National Institute of Allergy and Infectious Diseases to D.A.S. D.A.S. is also supported in part by funds from the National Cancer Institute (1PO1-CA-14236-05). We thank I. Sures for instruction in DNA sequencing techniques, B. Eaton for technical assistance, and B. Moss for providing samples of the vaccinia capping enzymes. We are grateful to L. Korn and S. H. Kim for instruction in computer analysis of nucleic acid sequences.

REFERENCES

1. Erdmann, V.A. (1979) Nucleic Acids Res. 6, r29-r44.
2. Hori, H. and Osawa, S. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 381-385.
3. Ford, P.J. and Southern, E.M. (1973) Nature New Biol. 241, 7-12.
4. The abbreviations used are: 5S rRNA, 5S ribosomal RNA; 5S rDNA, DNA encoding 5S rRNA; bp, base pair; kb, kilobase. Restriction endonucleases are given in the nomenclature proposed by Smith, H.O. and Nathans, D. (1973) J. Mol. Biol. 81, 419-423.
5. Rubin, G.M. and Sulston, J.E. (1973) J. Mol. Biol. 79, 521-530.
6. Procunier, J.D. and Tartof, K.D. (1975) Genetics 81, 515-523.
7. Brown, D.D., Wensink, P.C. and Jordan, E. (1971) Proc. Natl. Acad. Sci. U.S.A. 68, 3175-3179.
8. Brown, D.D. and Sugimoto, K. (1973) Cold Spring Harbor Symp. Quant. Biol. 38, 501-505.
9. Maizels, N. (1976) Cell 9, 431-438.
10. Cockburn, A.F., Newkirk, M.J. and Firtel, R.A. (1976) Cell 9, 605-613.
11. Bell, G.I., De Gennaro, L.J., Gelfand, D.H., Bishop, R.J., Valenzuela, P. and Rutter, W.J. (1977) J. Biol. Chem. 252, 8118-8125.
12. Ford, P.J. and Brown, R.D. (1976) Cell 8, 485-493.
13. Birkenmeier, E.H., Brown, D.D. and Jordan, E. (1978) Cell 15, 1077-1086.
14. Wu, G.J. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 2175-2179.
15. Ng, S.Y., Parker, C.S. and Roeder, R.G. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 136-140.

16. Weil, P.A., Segall, J., Harris, B., Ng, S.Y. and Roeder, R.G. (1979) J. Biol. Chem. 254, 6163-6173.
17. Nijhawan, P. and Marzluff, W.F. (1979) Biochemistry 18, 1353-1360.
18. Lu, A.L., Blin, N. and Stafford, D.W. in preparation.
19. Bickle, T.A., Pirrotta, V. and Imber, R. (1977) Nucleic Acids Res. 4, 2561-2572.
20. Modrich, P. and Zabel, D. (1976) J. Biol. Chem. 251, 5866-5874.
21. Blin, N., and Stafford, D.W. (1976) Nucleic Acids Res. 3, 2303-2308.
22. Chang, A.C.Y. and Cohen, S.N. (1978) J. Bacteriol. 134, 1141-1156.
23. Cohen, S.N., Chang, A.C.Y. and Hsu, L. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 2110-2114.
24. Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 3961-3965.
25. Guerry, P., LeBlanc, D.J. and Falkow, S. (1973) J.Bacteriol. 116, 1064-1066.
26. McDonnell, M.W., Simon, M.N. and Studier, F.W. (1977) J. Mol. Biol. 110, 119-146.
27. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560-564.
28. Laskey, R.A. and Mills, A.D. (1975) Eur. J. Biochem. 56, 335-341.
29. Cavanaugh, G.M., Ed. (1964) Formulae and Methods of the Marine Biological Chemical Room V, Woods Hole, Mass.
30. Palmiter, R.D. (1974) Biochem. 13, 3606-3615.
31. Barrell, B.G. (1971) in Procedures in Nucleic Acid Research, Cantoni, G.L. and Davies, D.R., Eds. Vol. 2, pp. 751-779 Harper and Row, New York.
32. Nishimura, S. (1972) Prog. Nucleic Acid Res. Mol. Biol. 12, 49-85.
33. Peattie, D.A. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 1760-1764.
34. Ensinger, M.J., Martin, S.A., Paoletti, E. and Moss, B. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 2525-2529.
35. Moss, B. (1977) Biochem. Biophys. Res. Commun. 74, 374-383.
36. Korn, L.J., Queen, C.L. and Wegman, M.N. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4401-4405.
37. Celma, M.L., Pan, J. and Weissman, S.M. (1977) J. Biol. Chem. 252, 9032-9042.
38. Tshudi, C. and Pirrotta, V. (1979) Proc. Natl. Acad. Sci. U.S.A. in press.
39. Valenzuela, P., Bell, G.I., Masiarz, F.R., De Gennaro, L.J. and Rutter, W.J. (1977) Nature 267, 641-643.
40. Maxam, A.M., Tizard, R., Skryabin, K.G. and Gilbert, W. (1977) Nature 267, 643-645.
41. Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V. and Brown, D.D. (1978) Cell 13, 717-725.
42. Korn, L.J. and Brown, D.D. (1978) Cell 15, 1145-1156.
43 Gilbert, W. (1976) in RNA Polymerase, Losick, R. and Chamberlin, M. Eds., pp. 193-205, Cold Spring Harbor Laboratory, New York.