In Vitro Transcription of Cloned 5S RNA Genes of the Newt *Notophthalmus*

BRIAN K. KAY, OTTO SCHMIDT, and JOSEPH G. GALL

Department of Biology and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511. Dr. Kay's present address is the Laboratory of Biochemistry, National Cancer Institute, Bethesda, Maryland 20205.

ABSTRACT Recombinant plasmids that carried genes coding for 5S ribosomal RNA of the newt, *Notophthalmus viridescens*, were transcribed in vitro with extracts of *Xenopus laevis* oocyte nuclei. Plasmids containing multiple repeats of the 5S gene and spacer directed accurate transcription of 5S RNA (120 bases). Individual repeat units were recloned by inserting Sau 3A restriction fragments into the Bam HI site of plasmid pBR322. Because each repeat was cut by the enzyme within the coding region, the inserts had incomplete coding regions at their ends and spacer sequences in the middle. The DNA of these subclones directed synthesis of a 5S-size RNA that contained both plasmid and 5S RNA sequences. Transcription initiated in the vector, proceeded through the gene segment coding for nucleotides 41–120, and terminated at the end of the gene. The initiation of in vitro transcription required neither the original 5′ flanking sequences of the spacer nor the first third of the gene. We conclude that intragenic DNA sequences control the initiation of transcription. Other subclones that include pseudogenes gave rise to some transcripts 156 nucleotides long. These long transcripts represented continuation of transcription through the 36-base-pair pseudogene that is located immediately downstream from the 5S gene. However, most transcripts of these subclones terminated at the end of the normal gene before the beginning of the pseudogene. It is probable that a run of four or more Ts serves as part of the termination signal.

Recent studies on DNA sequences that control transcription in eukaryotes have focused on genes transcribed by RNA polymerase III. These genes are short and in several cases their sequences and the sequences of their RNA products are known. Analysis has been facilitated by the development of in vitro transcription systems that use cell-free extracts as a source of RNA polymerase III and necessary cofactors (2, 26, 43). Selective transcription of cloned eukaryotic tRNA genes from yeast (27), *Drosophila* (31), *Bombyx* (14, 16), and *Xenopus* (22) has been demonstrated when DNA templates were incubated with *Xenopus laevis* oocyte nuclear extracts. Other genes transcribed by polymerase III, such as the virus-associated (VA) RNA genes of adenovirus (2, 40, 43, 44) and the 5S RNA genes of *Xenopus* (2, 21, 26), have also been accurately expressed in vitro. A simple assay of transcriptional activity and the ability to rearrange DNA segments in vitro permit identification of sequences that control initiation, termination, and processing events.

We have studied in vitro transcription of the 5S RNA genes of the newt *Notophthalmus viridescens*. Characterization of these genes (Kay and Gall, manuscript in preparation) revealed that there are two types of 5S DNA repeat units: a major type that is 231 base pairs (bp) long and contains a 120-bp coding region and a 111-bp spacer, and a minor species that is 269 bp long and is composed of a normal 120-bp coding region, a 36-bp pseudogene, and a 113-bp spacer. The 36-bp pseudogene is a repeat of the terminal third of the gene. With this information in mind we wanted to learn whether the genes in multiple repeat units were transcribed autonomously in vitro and whether the pseudogenes were transcribed at all.

We also wanted to look for the type of regulation of transcription recently reported by D. Brown and his collaborators for the *Xenopus* 5S RNA genes (3, 5, 30). They have elegantly demonstrated the existence of a short intragenic region that controls transcription initiation. Moreover, R. Roeder and colleagues (12) have purified a protein that binds specifically to the middle of the gene and is essential for transcription.

In this paper we report the selective, accurate transcription

THE JOURNAL OF CELL BIOLOGY · VOLUME 90 AUGUST 1981 323-331

of the cloned newt 5S RNA genes, using oocyte nuclear extracts of *Xenopus* and *Notophthalmus*. We find that the 5′ flanking spacer sequence and the first 30–40 bp of the gene can be removed without disrupting synthesis of a 5S-size RNA product. We have shown that this RNA contains both plasmid and gene sequences. The results of the transcription experiments suggest that control signals for initiation reside in the middle third of the 5S gene, as in *Xenopus*.

MATERIALS AND METHODS

Construction of Subclones of p5S101

Subclones were constructed by inserting Sau 3A fragments of the plasmid p5S101 into the Bam HI site of pBR322. The recombinant p5S101 was originally isolated and characterized by Kay and Gall (manuscript in preparation). Sau 3A was a gift from C. Yen, Yale University. Bam HI was purchased from Boehringer Mannheim Biochemicals, Indianapolis, Ind. The plasmid p5S101 was digested to completion with Sau 3A at 37°C in 60 mM NaCl, 7 mM Tris-HCl (pH 7.4), 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol. The fragments were resolved on a 2% low-melting agarose gel (SeaPlaque agarose; Seakem, Marine Colloids Div., Rockland, Maine) in a Bio-Rad model 220 vertical slab apparatus (Bio-Rad Laboratories, Richmond, Calif.) by electrophoresis in 40 mM Tris-acetate (pH 8.0), 2 mM EDTA, 5 mM Na-acetate. After electrophoresis, the gel was stained with ethidium bromide (0.5 μg/ml) for 30 min before being viewed with a UV Products C-50 transilluminator (Ultra-Violet Products, San Gabriel, Calif.). The visualized DNA bands 221 and 259 bp long were excised and melted at 65°C with an equal volume of water. Before phenol extraction, 0.1 volume of 3 M Na-acetate was added to the liquified material. The aqueous phase was reextracted with phenol to ensure complete removal of agarose. Traces of phenol were removed by three extractions with ether, followed by bubbling air through the solution. The DNA was recovered by ethanol precipitation in the presence of 10 μg of carrier *E. coli* tRNA (Sigma Chemical Co., St. Louis, Mo.). The purified fragments were ligated separately to an inactivated Bam HI digest of pBR322 DNA (4). The ligations were done in 90 mM Tris-HCl (pH 7.6), 10 mM dithiothreitol, 1 mM ATP with 0.1 U of T4 DNA ligase (New England Biolabs, Beverly, Mass.) in a final volume of 50 μl. After an overnight incubation at 5°C, the ligation mixture was heat inactivated and restricted with Bam HI. The enzyme linearized only nonrecombinant plasmids because the insertion of the Sau 3A fragments destroys the Bam HI site; the Sau 3A sites in the 5S DNA are flanked on both sides by T residues. Transformation was carried out according to Wensink et al. (42) with *E. coli* HB101. Ampicillin-resistant and tetracycline-sensitive colonies were picked and their recombinant plasmids were isolated. Growth of clones, amplification of plasmids, and isolation of circular plasmid DNA was according to the methods of Tanaka et al. (36) and Clewell and Helenski (8). All work with recombinant DNA was performed under P2 and EK1 containment conditions according to the 1976 National Institutes of Health Guidelines for recombinant DNA research.

Transcription Reactions and Product Analysis

Extracts of oocyte nuclei from *Xenopus laevis* and *Notophthalmus viridescens* were prepared according to the method of Birkenmeier et al. (2). *Xenopus* and *Notophthalmus* females were obtained from the South African Snake Farm and Lee's Newt Farm, respectively. The oocytes of *Notophthalmus* were not primed, because their yolk has a consistency that does not require priming. Transcripts were electrophoresed in 10% polyacrylamide/4 M urea gels that were 40 cm × 20 cm × 0.3 cm with 90 mM Tris-borate (pH 8.5), 3 mM EDTA buffer (24). After radioautography, the appropriate gel slices were eluted electrophoretically through a small column of polymerized polyacrylamide, and the RNA was recovered by ethanol precipitation. Digestion of the [^{32}P]RNA with RNase T_1 (Sankyo, Tokyo) and subsequent separation of the oligonucleotides by two-dimensional ionophoresis was performed according to Barrell (1). The first dimension was on cellulose acetate strips at pH 3.5 and the second dimension was on DEAE paper at pH 1.7. After radioautography, spots of paper were cut out and the radioactive oligonucleotides were eluted for secondary analysis. RNase A (Worthington Biochemical Corp., Freehold, N. J.) digests were electrophoresed on DEAE paper at pH 3.5 for structure determination. RNase T_2 (Sankyo) digests were chromatographed on polyethyleneimine-cellulose plates according to Goody and Eckstein (15).

Isolation of 5′-[γ-S]RNA on mercury-agarose was performed according to Silverman et al. (32). 5′-[γ-S]GTP was purchased from Boehringer Mannheim Biochemicals and mercury-agarose (Affi-Gel 501) was obtained from Bio-Rad Laboratories.

RESULTS

Transcription of Cloned Newt 5S RNA Genes

The recombinant plasmids p5S101 and p5S102 contain, respectively, 4.3 and 3.5 × 10^3 bp of *N. viridescens* DNA cloned in the Bam HI site of plasmid pBR322 (Kay and Gall, manuscript in preparation). Restriction enzyme maps show that p5S101 and p5S102 are made up of 18 and 15 tandem 5S DNA repeat units. The repeats in p5S102 are homogenous in length, 231 bp long. In p5S101 there are 14 of the 231-bp units interspersed with four 269-bp repeats. The longer repeats are identical in DNA sequence to the smaller ones except for a tandem duplication of the terminal third of the coding region and two nucleotides of the 3′ flanking spacer (Fig. 1). The duplicated coding segment is termed a pseudogene. The gene sequence has been identified by comparison with *Xenopus* 5S genes (21) and by preliminary sequencing of newt 5S RNA (data not shown).

p5S101 and p5S102 DNAs were used as templates for in vitro transcription with *Xenopus* oocyte nuclear extracts. Fig. 2 shows a radioautogram of the RNA synthesized. Both templates direct the synthesis of a discrete RNA that migrates on the gel in the region expected for 5S RNA (lanes *a* and *b*, RNA-1). The synthesis of this RNA requires exogenous cloned 5S DNA. Because a low concentration of α-amanitin (0.5 μg/ml) did not inhibit synthesis, but a high concentration (200 μg/ml) did (data not shown), it may be concluded that RNA-1 is transcribed by RNA polymerase III (41). RNA-1 transcribed from both templates was analyzed by RNase T_1 digestion and yielded identical fingerprints. The RNase T_1 fingerprint for RNA-1 synthesized by p5S101 is shown in Fig. 5 *a*. Secondary analysis of the T_1 oligonucleotides is consistent with the known sequence of the 5S genes in the plasmid with one exception (Table I). The exception is the dinucleotide CG in the fingerprint, whereas there is no GCG in the noncoding strand of the predominant 5S gene sequence (Kay and Gall, manuscript in preparation). Some minor DNA sequence heterogeneity does exist between repeats as evidenced by the presence of Hha I sites (GCGC) in two of the genes of p5S101. If these variant genes were being transcribed, they would yield the CG dinucleotide.

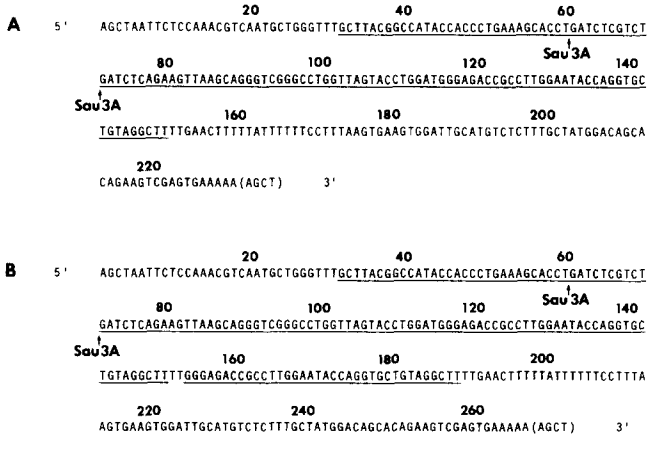

FIGURE 1 DNA sequences from the noncoding strand of the 231- and 269-bp (base pair) 5S DNA repeat units of the newt *Notophthalmus* (Kay and Gall, manuscript in preparation). In the 231-bp unit (*A*) the sequence of the gene is underlined; in the 269-bp unit (*B*) the gene and following pseudogene are underlined. The Sau 3A cleavage sites (5′ GATC 3′) are noted in the two DNA sequences.

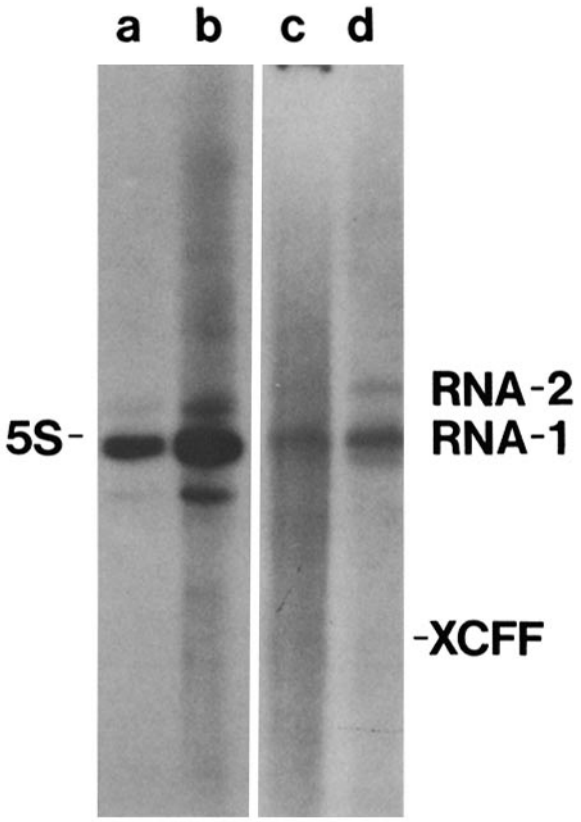

FIGURE 2 Gel electrophoretic analysis of RNA made in vitro by *Xenopus* oocyte nuclear extracts and cloned newt 5S RNA genes. Transcripts were prepared with [α-^{32}P]GTP from two recombinant plasmids, p5S101 (lane *a*) and p5S102 (lane *b*), which contain 18 and 15 tandem 5S DNA repeats respectively. These DNAs direct synthesis of two different sized RNAs, RNA-1, which migrates to the same position as 5S RNA, and RNA-2, which migrates as if it is ~10 nucleotides longer. To demonstrate that these two RNAs have 5′-triphosphate termini, indicative of primary transcripts, RNA was synthesized from p5S101 DNA with 5′-[γ-S]GTP and [α-^{32}P]UTP, fractionated on a mercury-agarose column, and then electrophoresed on a polyacrylamide gel. RNA that did not contain γ-S and was not retained on the column is shown in lane *c*. RNA that contained γ-S and was retained on the column is shown in lane *d*. Both RNA-1 and RNA-2 have 5′-triphosphate termini. The gel was 10% polyacrylamide with 4 M urea. The positions of xylene cyanol FF dye (*XCFF*) and authentic 5S RNA are noted on the radioautogram.

The minor transcription product, RNA-2, migrates slightly slower than RNA-1 (Fig. 2, lanes *a* and *b*). It has not been characterized because of a scarcity of radioactive material. However, as data presented below will show, RNA of this size results from transcription beyond the 3′ end of the 5S gene. RNA-2 migrates as if it were ~10 nucleotides larger than RNA-1.

To ascertain whether the synthesized RNAs have unprocessed 5′ ends characteristic of primary transcripts, in vitro transcription of p5S101 was repeated in the presence of 5′-[γ-S]GTP. Because the sulfur group is located on the 5′ terminal phosphate only RNA with a polyphosphate 5′ terminus will carry this group (15). The substrate did not interfere with transcription and the resulting products were chromatographed over a mercury-agarose column to separate RNAs containing γ-S triphosphate from those that did not (32). Transcripts of p5S101 fractionated in this manner and resolved by gel electrophoresis are shown in Fig. 2 (lanes *c* and *d*). It is evident that both RNA-1 and RNA-2 carry a 5′ triphosphate group (lane *d*). Not all transcripts this size, however, were bound to the column (lane *c*). These transcripts are probably RNAs that had been initiated with a normal guanosine triphosphate, either endogenously present in the nuclear extract or the result of exchange of the γ-S from [γ-S]GTP to the gamma position of other nucleoside triphosphates (20).

To demonstrate specific initiation in vitro we identified the 5′ terminal nucleotides of the RNA transcribed from p5S101 in two ways. First, when this RNA is synthesized in the presence of [α-^{32}P]GTP and cleaved by RNase T_1, there is a fragment that chromatographs as pppGp on polyethyleneimine-cellulose (PEI) plates (Fig. 5 *a*, pppG). Second, when the substrate is [α-^{32}P]CTP there is a nearest neighbor transfer of label to pppGp after RNase T_2 digestion of the 5S RNA (data not shown). These data demonstrate that the first two nucleotides of the transcript are pppGC. Because the first two nucleotides of the 5S gene are also GC (Fig. 1), it appears that transcription of the cloned newt genes by *Xenopus* oocyte nuclear extracts occurs in a specific manner, with accurate initiation.

Transcription of p5S101 Subclones

From the analysis presented above it was not known how many repeats in p5S101 and p5S102 were active templates for 5S RNA synthesis. To evaluate the coding ability of single genes as well as the pseudogene, we constructed subclones that carried one repeat unit per plasmid. The subclones were constructed by inserting Sau 3A restriction fragments from p5S101 into the Bam HI site of plasmid pBR322 (see Materials and Methods). Sau 3A cuts twice in both types of 5S repeat units present in p5S101, at nucleotide positions 30/31 and 40/41 of the coding region. None of the subcloned units contains Hha I sites. In the subclones diagrammed in Fig. 3 the 5S gene is split into two segments: an anterior segment coding for nucleotides 1–30 (short box), and two kinds of posterior segment. One posterior segment codes for nucleotides 41–120 (long box), whereas the other codes for the same region plus the pseudogene (nucleotides 85–120, hatched box). Because of this arrangement, the original 5′ flanking spacer sequence and the first third of the gene are now separated from the bulk of the gene. For two of the subclones, nucleotide 41 is adjacent to vector DNA sequences. The other two subclones carry the small 10-bp Sau 3A restriction fragment coding for nucleotides 31–40 attached to the posterior gene segment. Pairs of subclones (1 and 3, 2 and 4) carry the 5S DNA repeat units in opposite orientations in the vector. We have been able to demonstrate that the rearranged genes of these subclones are transcriptionally functional and that the control of transcription initiation is probably the same as recently shown for *Xenopus* 5S genes by Brown and co-workers (3, 30).

Fig. 4 is a radioautogram of RNA transcribed from the four subclones by *Xenopus* oocyte nuclear extracts. Even though the 5S gene is split into two separate pieces, each recombinant is capable of synthesizing 5S-size RNA, plus some larger products (lanes *b–e*). The electrophoresis patterns are the same for the pairs of inserts in opposite orientations in the vector (lanes *b* and *d*, *c* and *e*).

TABLE I

RNase T1 End Products of 5S-size RNA

	p5S101 transcript		Subclone 1 transcript			
Designation	GTP§	Deduced*	ATP§	UTP§	GTP§	Deduced‡
T1	U	CCAUACCACCUG				
T2	AG	AUCUCAG	AG, C	AU, C	AG	AUCUCAG
T3	AC, U, C, G	AUCUCG CUUACG UACCUG	U	ND	G, U	UACCUG(G)
T4	AAG	UUAAG	AAG, U	U	AAG	UUAAG
T5	AG	UUAG	U	AG, U	AG	UUAG(U)
T6	U	UCUG				
T7	AG	AAUACCAG	AAU, C	AAU	AG	AAUACCAG
T8	U	CACCUG				
T9	G, U	CCUG(G)		C	G, U	CCUG(G)
T10	AU, G	AUG(G)		AU	AU, G	AUG(G)
T11	AG	UAG	U		AG	UAG
T12	G, C, U	CUG UCG		G, C	G, C, U	CUG UCG
T13	G, U	UG (G)			G, U	UG(G)
T14	AAAG	AAAG				
T15	C	ACCG			C	ACCG
T16	AAG	AAG	AAG	AAG	AAG	AAG(U)
T17	AG	CAG	C		AG	CAG
T18	AG	AG	AG		AG	AG
T19	C	CG	G		C	CG(A)
T20	G	G	G	G	G	
A			Origin‖, G	Origin‖	ND	pppAUCG(A)
B			ND	G	ND	ACCACACCCG(U)
C			C	AU	AU, G	AUCAUG(G)
D				G, C	U	UCCUG(U)
E			U	AC	AC	ACUACG

RNase T_1 end products of 5S-sized RNA. The RNase T_1 oligonucleotides of labeled RNA synthesized by p5S101 and subclone 1 were analyzed secondarily with RNase A and RNase T_2. p5S101 transcripts made with radioactive GTP were cleaved with RNase T_1 to produce oligonucleotides designated T. These oligonucleotides were subsequently cleaved with RNase A and the labeled products are listed. From this information and the DNA sequence, a sequence was deduced which matches the gene sequence. This same type of analysis was likewise done on subclone 1 transcripts that were labeled separately with radioactive ATP, UTP, and GTP. These data show that the two transcripts, although similar, do differ in sequence. This difference results from the transcription of pBR322 adjacent to the gene segment 41–120 in subclone 1.

* Sequence deduced on basis of mobility and labeling data.

‡ Bases indicated in parentheses were deduced from nearest neighbor analysis.

§ Digested with RNase A.

‖ RNase T_2 digestion and separation on PEI revealed pppAp.

Sequence Analysis of Subclone 5S-size Transcripts

To establish the identity of the transcription products, we eluted the bands of labeled RNA from the gel and analyzed the nucleotide sequence by conventional methods (1). Fig. 5 compares the RNase T_1 fingerprints of 5S-size RNA synthesized by p5S101 (*a*) and subclone 1 (*c*) in the presence of [α-^{32}P]GTP. Fig. 5 *b* is a summation of data from both fingerprints. The numbered oligonucleotides designate those present in the p5S101 transcript, some of which are shared by the subclone 1 transcript, whereas the lettered oligonucleotides refer to those unique to the subclone 1 transcript. To complete the sequence analysis additional transcripts were prepared, using different [α-^{32}P]NTPs. The secondary analysis of the T_1 oligonucleotides is given in Table I.

Knowing the sequence of pBR322 (34, 35) and of the newt 5S DNA, and using the sequence analysis in Table I, we deduced the structure of the 5S-size transcript of subclone 1. This transcript begins in the pBR322 DNA adjacent to the larger 5S coding segment and extends to the 3′ end of the gene (Fig. 3). Oligonucleotides that match the corresponding pBR322 DNA sequence are indicated in Fig. 6. All oligonucleotides unique to transcripts of subclone 1 fall within the stretch of vector DNA. From the sequence analysis we conclude that neither the 5′ flanking spacer nor the first 40 bp of the gene are necessary for transcription of the newt 5S RNA gene.

We determined that the 5′ and 3′ ends of the 5S-sized transcript of subclone 1 are pppAU and CUU_{OH}. Fragment A (Fig. 5 *c*) released pppAp after RNase T_2 digestion, as identified by chromatography on PEI-plates. Because this nucleoside tetraphosphate is labeled in RNase T_2 digests of transcripts synthesized in the presence of either [α-^{32}P]ATP or [α-^{32}P]UTP, the first two nucleotides must be A and U. Because the oligonucleotide that includes the dinucleotide AU has an A residue as the nearest neighbor of a G (Table I), we conclude that initiation for subclone 1 most likely occurs at position 3 (Fig. 6). The expected start site, a C at position 1 (assuming that transcription compensates exactly for the gene's missing 40 bp), is not used. Because transcription proceeded from pBR322 into the terminal part of the gene, it was expected that the transcript would terminate normally. This prediction was substantiated by the demonstration of a fragment that behaved as CUU_{OH} (39) in RNase T_1 digests of transcripts synthesized in the presence of [α-^{32}P]UTP (data not shown). Taken together, our data indicate that the 5S-size transcript of subclone 1 is 118

nucleotides long. Because the 5S-size RNA synthesized by subclone 2 has the same T_1 fingerprint as the subclone 1 transcript, we assume them to be identical.

The exact initiation site for subclones 3 and 4 is not known, because their RNA-1 products were not sized precisely nor sequenced. We suggest a probable site (position 4 in Fig. 6) based on two observations: first, RNA-1 from the two subclones appears to be smaller than 120 nucleotides (Fig. 4) and second, RNA polymerase III has a known preference for initiating at purines that are preceded and followed by pyrimidines (20).

Sequence Analysis of Subclone Transcripts Larger Than 5S Size

Fingerprinting and secondary analysis of RNA-2 transcribed from subclone 1 confirm that it is an extended version of RNA-1, with the additional nucleotides at the 3′ end. The RNase T_1 fingerprints for RNA-1 and RNA-2 of subclone 1, synthesized in the presence of [α-^{32}P]ATP, were identical except for two extra spots in the RNA-2 fingerprint (data not shown). The 3′ end of RNA-1, which is CUU_{OH}, was not labeled in this experiment. The supplementary oligonucleotides, identified as CUUUUG and $AACUUUU_{OH}$, are coded for by sequences of the gene and the spacer (Fig. 7). The finding of 10 additional nucleotides at the end of RNA-2 agrees with the length estimate based on gel electrophoresis. Transcripts extended at the 3′ end have been observed during in vitro transcription of *Xenopus* 5S RNA genes (2, 40).

Subclones 2 and 4, which carry the minor 5S repeat with a pseudogene, synthesize two size classes of RNA that are much larger than 5S (Fig. 4, lanes *c* and *e*). RNA-3 and RNA-4 migrate on gels as if they are approximately 160 and 170 nucleotides long, respectively. The fingerprints of RNA-1 and RNA-3 synthesized by subclone 2 in the presence of [α-

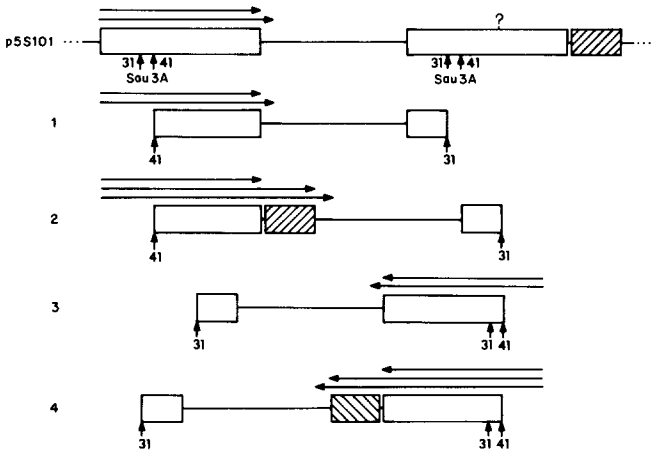

FIGURE 3 The structure and transcription schemes of p5S101 and subclones. The 5S DNA repeat units are displayed with boxes and lines representing gene and spacer sequences respectively. The hatched boxes correspond to the 36-bp pseudogene. Two repeat units of p5S101 are shown with their Sau 3A cleavage sites. The four subclones were constructed by insertion of Sau 3A fragments into the Bam HI site of pBR322. In p5S101 and the four subclones, the Sau 3A cleavage sites at positions 31 and 41 in the gene are designated. The major in vitro transcripts of these five DNA templates are represented by arrows that show the DNA sequences transcribed and the direction of transcription (see text).

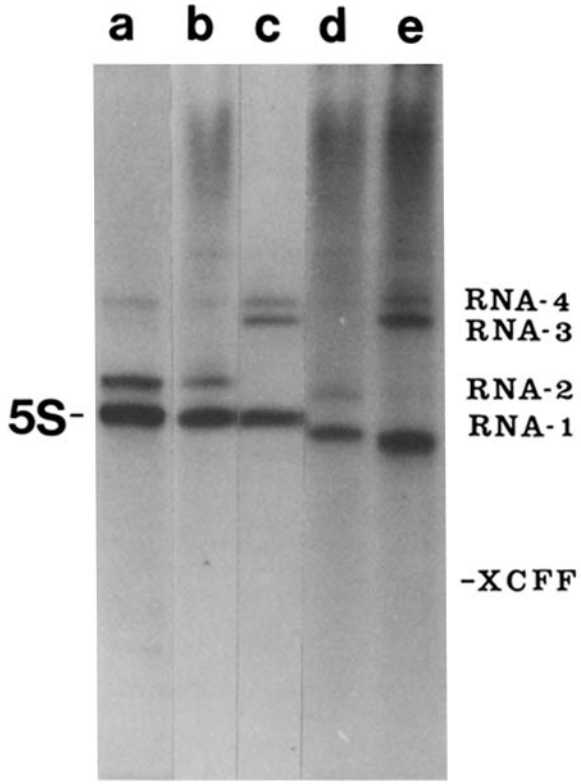

FIGURE 4 Gel electrophoretic analysis of RNA transcribed in vitro by p5S101 and subclones. Transcripts were prepared from the following DNA templates: (*a*) p5S101, (*b*) subclone 1, (*c*) subclone 2, (*d*) subclone 3, and (*e*) subclone 4. The subclones contain single 5S gene repeats whose coding regions were split into two segments by the cloning procedure (Fig. 3). In each case a 5S-size RNA is synthesized (RNA-1); that from p5S101 is authentic 5S RNA, whereas that from the subclones contains some vector and some 5S RNA sequences. Larger transcripts (RNA-2, 3, 4) result from readthrough beyond the end of the 5S gene, as described in the text. The gel was 10% polyacrylamide and contained 4 M urea. *XCFF* is the position of xylene cyanol FF dye.

^{32}P]GTP are displayed in Fig. 8. These two patterns match except for one additional fragment in the RNA-3 fingerprint. This fragment (denoted by an arrow in Fig. 8 *a*), corresponds to CUUUG. The remaining 35 or so nucleotides of the 160-nucleotide-long RNA-3 do not produce uniquely migrating oligonucleotides. The simplest explanation is that RNA-3 is a 156-nucleotide product that initiated in the vector, transcribed through the posterior gene segment and the trailing pseudogene, and terminated at the end of the pseudogene (Fig. 3, arrows). Such transcription would generate only one new oligonucleotide, CUUUUG, corresponding to sequences between the end of the gene and the beginning of the pseudogene. Although the largest transcript of subclone 2 has not been characterized, it is likely that it is an extended version of RNA-3, continuing to a second termination site in the spacer (Fig. 7). We have not ruled out other possibilities.

The RNA transcripts of subclones 3 and 4 have not been sequenced. Because the products of these subclones have mobilities comparable to those from subclones 1 and 2, it is probable that the orientation of the subclone in the vector is not important for transcription. Fig. 3 summarizes the probable transcription schemes for the four subclones.

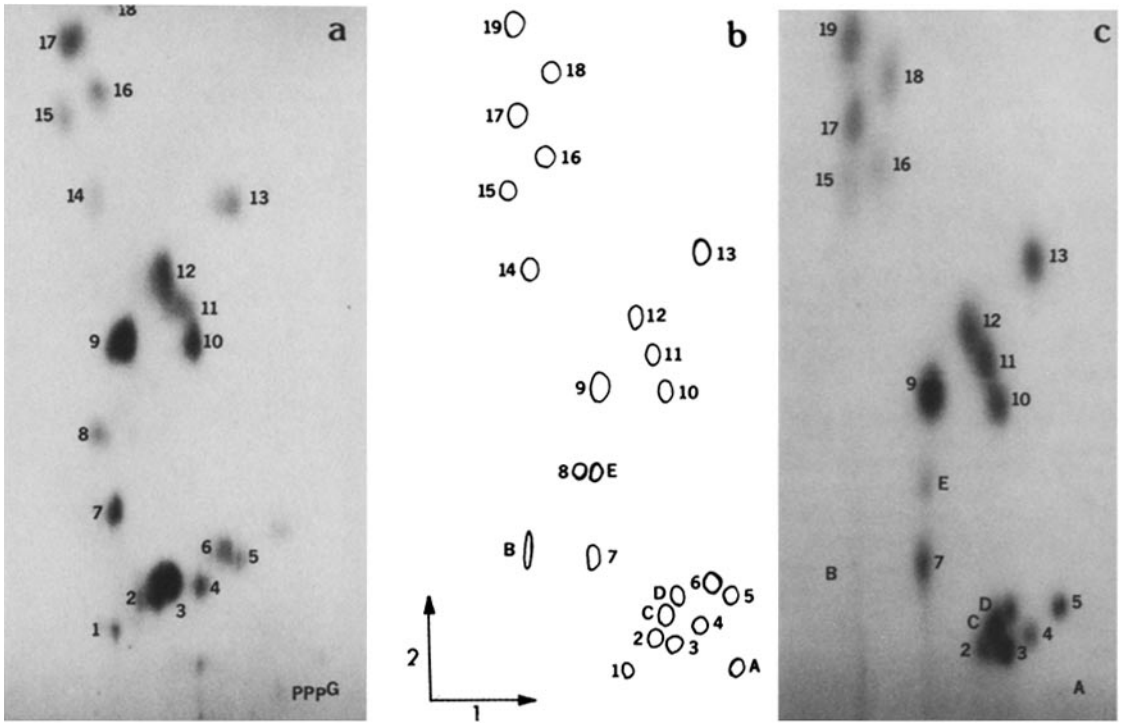

FIGURE 5 Fingerprint analysis of 5S-sized in vitro transcripts of p5S101 DNA (*a*) and subclone 1 DNA (*c*). RNA-1, transcribed from these two DNAs in the presence of [α-^{32}P]GTP, was cleaved with RNase T_1 and the resulting oligonucleotides were fractionated by two-dimensional electrophoresis. The oligonucleotides in panel *a* are numbered; some of these are present in panel *c*. The oligonucleotides in panel *c* that do not correspond to 5S RNA sequences are lettered. The data, summarized in panel *b*, show that the 5S-size transcript of subclone 1 contains both 5S RNA and pBR322 sequences. The arrows denote the two dimensions of separation; the first dimension was on cellulose acetate strips at pH 3.5 and the second dimension was on DEAE paper at pH 1.7. Sequence analysis of the two RNAs is further described in Table I and Fig. 6.

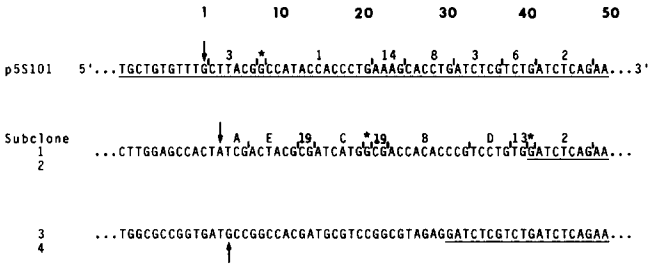

FIGURE 6 Nucleotide sequences surrounding the transcription initiation sites of p5S101 and the four subclones. The noncoding strand is shown for part of the 5S DNA in p5S101, subclones 1 and 2, and subclones 3 and 4. Sequences corresponding to the 5S gene and spacer are underlined, whereas the pBR322 sequences (35) are not. RNase T_1 oligonucleotides of in vitro transcripts, depicted in Fig. 5 and described in Table I by numbers and letters, are indicated above the corresponding DNA sequences. Numbered oligonucleotides correspond to sequences present in the 5S RNA made by p5S101. Lettered oligonucleotides correspond to pBR322 sequences present in the 5S-size transcripts of subclones 1 and 2. The symbol * designates the RNase T_1 product T20. The 5S-size transcripts of subclones 3 and 4 were not sequenced. At the top of the figure, the first 50 nucleotide positions of the 5′ end of the 5S gene are noted. Arrows marked initiation sites (i.e., first nucleotide of transcripts); the downward arrows mark experimentally determined sites, whereas the upward arrow marks the probable site (20).

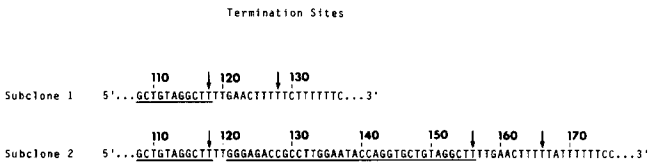

FIGURE 7 Nucleotide sequences that surround the in vitro termination sites of subclones 1 and 2. Sequences are shown for the noncoding strand, with gene and pseudogene sequences underlined. The numbers above the sequences denote nucleotide positions downstream from the initiation sites. The arrows point to the major termination sites used during in vitro transcription experiments (see text). Note that readthrough of transcription in subclone 2 gives rise to a 156-nucleotide-long RNA that contains the diagnostic T_1 oligonucleotide CUUUUG (Fig. 8).

Transcription of Subclones with Notophthalmus Germinal Vesicle Extracts

We have tested the transcriptional activity of nuclear extracts prepared from *Notophthalmus* oocytes by use of newt 5S DNA recombinant plasmids as templates. Generally, the extracts did not transcribe as efficiently as did the *Xenopus* oocyte nuclear extracts. Fig. 9 is an example of in vitro transcription in the homologous system. The presence of 5S-size transcripts indicates that specific initiation and termination takes place on subclones 1, 2, and 3. The origin of the transcripts that are smaller than 5S size is not known, but similar minor transcripts were also occasionally seen with *Xenopus* extracts. Several attempts have been made to improve the transcriptional efficiency of the homologous system. So far neither nuclear extracts from oocytes at different stages of oogenesis nor different buffer conditions have resulted in higher yields of in vitro transcription.

DISCUSSION

Selective Transcription of Newt 5S RNA Genes

We have shown here that recombinant DNA plasmids carrying single and multiple repeats of the 5S RNA genes of the newt *Notophthalmus* can be transcribed accurately in vitro. The transcripts initiated with purine triphosphates and terminated in U-rich oligonucleotides, like other eukaryotic 5S RNAs (18). Because we used *Xenopus* oocyte nuclei as the source of RNA polymerase III and cofactors for transcription of *Notophthalmus* 5S genes, there may be some question whether the initiation and termination sequences are the ones used in the newt. However, the coding region of *Notophthalmus* 5S DNA shows 91% homology with the *Xenopus laevis* oocyte-type 5S gene and this degree of homology is not far outside the range observed among the four *Xenopus* 5S RNA multigene families that have been successfully transcribed in vitro with oocyte nuclear extracts of *Xenopus laevis*. Moreover, oocyte nuclear extracts of *Notophthalmus* direct the same qualitative pattern of transcription as do *Xenopus* extracts.

Intragenic Location of Initiation Control Signals

Transcription experiments with 5S DNA repeat units derived from plasmid p5S101 establish that the subclones are competent to direct synthesis of 5S-size RNA transcripts, even though the genes are broken into separate anterior and posterior segments. We found that the 5′ end of the transcripts is coded for by vector DNA and the 3′ end by the posterior segment of the gene (bases 41–120). We did not detect transcripts of the anterior segment (bases 1–40). Apparently, therefore, the information to control selective initiation of newt 5S RNA genes does not reside in either the 5′ end of the gene or the 5′ flanking sequences. That the subclones produced a 5S-size transcript, regardless of the orientation of the insert in the vector, implies

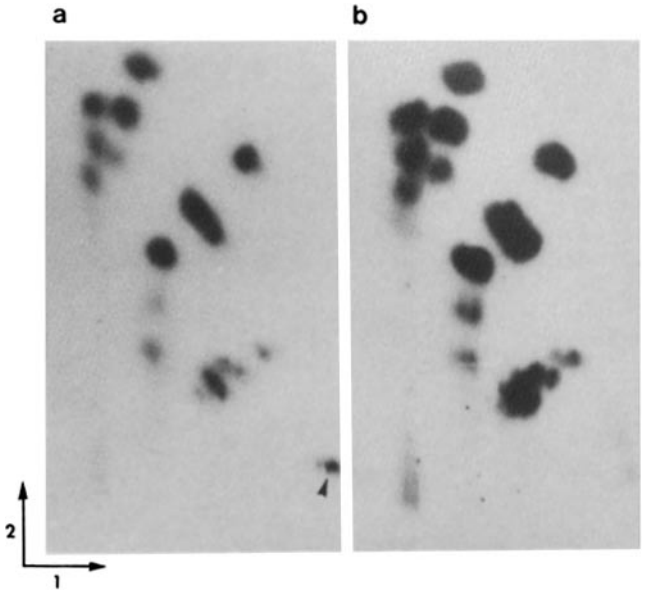

FIGURE 8 Fingerprint analysis of subclone 2 transcripts. Radioautograms are shown of the two-dimensional separation of the RNase T_1 products of RNA-2 (*a*) and RNA-1 (*b*) synthesized by subclone 2 DNA with [α-^{32}P]GTP. The arrows denote the first and second dimensions of electrophoresis (see Fig. 5). The patterns are identical except for the oligonucleotide CUUUUG, marked by a small arrowhead in *a*.

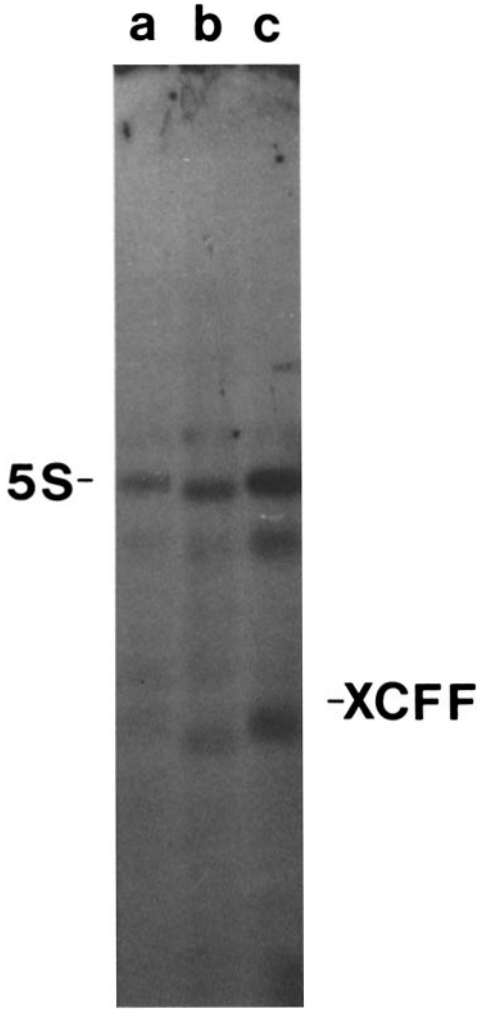

FIGURE 9 In vitro transcription using an extract from newt oocyte nuclei. A radioautogram is shown of the electrophoretic analysis of 5S-size RNA synthesized by newt extracts using as templates DNA of subclone 1 (*a*), subclone 2 (*b*), and subclone 3 (*c*). The gel was 10% polyacrylamide and contained 4 M urea. The position of XCFF and authentic 5S RNA are marked.

that initiation of RNA synthesis is not strongly influenced by the sequences that substitute for the first 30–40 nucleotides of the gene and the 5′ flanking spacer. These observations are consistent with the conclusion reached by D. Brown and coworkers (3, 30) that initiation is controlled by sequences in the middle of the coding region. Using synthetic deletions of *Xenopus borealis* somatic-type 5S DNA, these investigators found that the entire 5′ flanking sequence and the first 50 nucleotides of the gene could be deleted without disrupting synthesis of 5S-size RNA by *Xenopus laevis* oocyte nuclear extracts. In an experiment similar to that reported here (7), a recombinant plasmid that lacked the first 40 nucleotides of the *Xenopus borealis* somatic-type 5S gene was constructed by inserting a Sau 3A restriction fragment of 5S DNA into the Bam HI site of pBR322. This recombinant supported synthesis of 5S-size RNA.

The in vitro transcription results, with both newt and frog 5S genes, imply that initiation of transcription occurs at a fixed distance upstream from a reference point. The distance is not rigidly fixed, however, as shown by the fact that newt transcripts were either 118 or 120 nucleotides long. A range of 116

to 121 nucleotides was observed with the mutant *Xenopus* 5S genes (30). The choice of the exact initiation site may be influenced by the location of purines in the DNA, because purines are the preferred starting nucleotide of *Xenopus laevis* RNA polymerase III (20). The sequence surrounding the approximate site may also influence the choice, as demonstrated by the observation that certain 5′ flanking sequence deletions of *Xenopus* 5S genes (30) and adenovirus (Ad5) VA genes (38) have altered starting site selection.

We observed that the pseudogenes in the newt 5S subclones could be transcribed as part of longer transcripts. This "piggy-back" transcription further suggests that the DNA sequences that control initiation are in the middle of the gene. Suppose, for instance, that initiation began a fixed distance upstream from a control region in the DNA flanking the 3′ end of the gene. In this case, the pseudogene should cause initiation of transcription to be shifted downstream ~40 nucleotides from the initiation site actually observed. Sakonju et al. (30) found that several insertions within the *Xenopus* 5S gene caused the initiation site to be shifted downstream by distances equal to the length of the insert. In these "maxigenes" the insertions were upstream from the control region. On the other hand, if a control region were located within the pseudogene (nucleotides 85–120) there would be two control regions, one in the gene and one in the pseudogene. Such an arrangement might lead to conflicting initiation events. Because initiation does occur at the correct position even in the presence of the 3′ flanking pseudogene, the simplest interpretation is that the control region lies before nucleotide 85, the first nucleotide of the pseudogene. We realize that the data presented here are only suggestive and not conclusive concerning the location of the control region. All our observations are consistent with the hypothesis that the control region for transcription of the *Notophthalmus* 5S gene is the same as for the *Xenopus* gene, namely between nucleotides 50–83 (30).

DNA sequences that control initiation may be intragenic for other genes transcribed by RNA polymerase III. Cloned tRNA genes from *Bombyx* $tRNA_1^{Ala}$ (14) and *Xenopus* $tRNA_1^{Met}$ (37) with short 5′ flanking sequences are accurately expressed by the transcriptional machinery of *Xenopus* oocyte nuclei. Expression of the *Xenopus* $tRNA_1^{Met}$ gene requires, however, both the 5′ and 3′ ends of the cloned gene unit (22). Recently, manipulation of cloned *Drosophila* $tRNA^{Lys}$ genes has shown that these genes can be transcribed effectively by *Xenopus* oocyte nuclear extracts without the original 5′ flanking sequences (10, 31). In plasmids that lacked the normal 5′ sequences, transcription initiated in the vector approximately the same number of nucleotides upstream from the gene as did transcription of the original cloned gene. Finally, experiments on in vitro transcription of the adenovirus VA gene show that a synthetic deletion mutant missing 9 bp of the 5′ end of the gene can be transcribed faithfully (13).

Transcription Termination Signals

The products transcribed from subclones containing the newt pseudogene help define the DNA sequences that control termination of in vitro transcription. Although the subclones direct some synthesis of piggy-back transcripts, the predominant product is 5S-size RNA. From the sequence homology and the apparent termination sites shown in Fig. 7, it would seem likely that termination involves a cluster of T residues. The DNA sequences downstream from the cluster are probably not important because termination at the end of the 5S gene in subclones 1 and 2 occurs even though the sequences downstream are different. Two features, the presence of hyphenated dyad symmetry elements in the gene and clusters of T residues at the end of the gene, have been implicated as termination signals of prokaryotic and eukaryotic genes (21, 28, 33). These two features are present in the newt 5S RNA gene (i.e., in Fig. 1 *a*, a dyad symmetry element is present at residues 108–113 and 135–140, while the largest T clusters are at residues 150–153, 158–162, and 164–169). For proper termination of *Xenopus* 5S genes, a block of at least four T residues is required at the end of the gene (3). This requirement is exemplified by the inefficient termination of a *Xenopus laevis* 5S gene in a plasmid where there is a terminal cluster of three T residues instead of four (6, 7). The possible role of hyphenated dyad symmetry elements in the gene near the 3′ end has not been experimentally established.

Three T clusters four, five, and six residues long (Fig. 9) follow the coding region of the 231-bp newt 5S repeat. In vitro transcripts of p5S101 and its subclones terminate within either the four or five residue T clusters, without continuing to the third cluster that is only 13 nucleotides from the end of the gene. Apparently a cluster of four T residues is an adequate in vitro termination signal but a cluster of five is more effective. The presence of three potentially functional termination signals ensures complete termination at the end of gene without substantial transcription of the adjacent spacer. The correlation between termination effectiveness and T cluster length is supported by recent observations that show that rU-dA (RNA-DNA) hybrids are very unstable (25).

In the newt 269-bp 5S repeat unit, there are four T clusters: one between the gene and pseudogene, and three beyond the pseudogene. When transcription proceeds past the first T cluster, the pseudogene is transcribed in its entirety. Whether or not such piggy-back transcription occurs in vivo is not known. In cultured *Xenopus* oocytes, the presence of 5S transcripts extended at the 3′ end (11) suggests that readthrough past the end of the genes can occur in vivo. Longer than normal 5S transcripts have also been seen in heat-shocked *Drosophila* cells (19, 23, 29) and in isolated HeLa and mouse cell nuclei (17), suggesting that extended transcripts may be widespread among eukaryotes and may function as precursors in 5S RNA synthesis.

We thank D. Söll for generous use of his laboratory facilities and P. Rae for helpful criticisms.

This work was supported in part by NIH research grants to J. G. Gall and Dieter Söll. B. K. Kay was supported by an NIH predoctoral training grant and O. Schmidt by a postdoctoral fellowship from the Deutsche Forschungsgemeinshaft.

This paper is based on a dissertation submitted by B. K. Kay in partial fulfillment of the requirements for the PhD degree at Yale University.

Received for publication 3 December 1980, and in revised form 13 April 1981.

REFERENCES

1. Barrell, B. G. 1971. Fractionation and sequence analysis of radioactive nucleotides. *Prog. Nucleic Acid Res.* 2:751–779.
2. Birkenmeier, E. H., D. D. Brown, and E. Jordan. 1978. A nuclear extract of *Xenopus laevis* oocytes that accurately transcribes 5S RNA genes. *Cell.* 15:1077–1086.
3. Bogenhagen, D. F., S. Sakonju, and D. D. Brown. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3′ border of the region. *Cell.* 19:27–35.
4. Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose system. *Gene (Amst.).* 2:95–113.
5. Brown, D. D., E. H. Birkenmeier, D. Bogenhagen, L. J. Korn, E. Jordan, R. Peterson, and

S. Sakonju. 1979. A dual 5S RNA gene system in *Xenopus*. *Carnegie Inst. Wash. Year Book*. 78:71–84.
6. Brown, D. D., and J. B. Gurdon. 1978. Cloned single repeating units of 5S DNA direct accurate transcription of 5S RNA when injected into *Xenopus* oocytes. *Proc. Natl. Acad. Sci. U. S. A.* 75:2849–2853.
7. Brown, D. D., and E. Jordan. 1979. A termination mutant of 5S DNA. *Carnegie Inst. Wash. Year Book*. 78:81.
8. Clewell, D. B., and D. R. Helinski. 1970. Properties of a supercoiled deoxyribonucleic acid-protein relaxation complex and strand specificity of the relaxation event. Biochemistry 9:4428–4440.
9. Davidson, E. 1976. Gene Activity in Early Development. Academic Press, Inc., New York.
10. DeFranco, D., O. Schmidt, and D. Soll. 1980. Two control regions for eukaryotic tRNA gene transcription. *Proc. Natl. Acad. Sci. U. S. A.* 77:3365–3368.
11. Denis, H., and M. Wegnez. 1973. Recherches biochimiques sur l'oogenèse. 7. Synthèse et maturation du RNA 5S dans les petits oocytes de *Xenopus laevis*. *Biochimie (Paris)*. 55: 1137–1151.
12. Engelke, D. R., S.-Y. Ng, B. S. Shastry, and R. G. Roeder. 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*. 19: 717–728.
13. Fowlkes, D. M., and T. Shenk. 1980. Transcriptional control regions of the adenovirus VAI RNA gene. *Cell*. 22:405–414.
14. Garber, R. L., and L. P. Gage. 1979. Transcription of a cloned *Bombyx mori* $tRNA^{ala}$ gene: nucleotide sequence of the tRNA precursor and its processing *in vitro*. *Cell*. 18:817–828.
15. Goody, R. S., and F. Eckstein. 1971. Thiophosphate analogs of nucleoside di- and triphosphates. *J. Chem. Soc. (Lond.)*. 93:6252–6257.
16. Hagenbüchle, O., D. Larson, G. I. Hall, and K. Sprague. (1979). The primary transcription of a silkworm alanine tRNA gene: identification of *in vitro* sites of initiation, termination and processing. *Cell*. 18:1217–1229.
17. Hamada, H., M. Muramatsu, Y. Urano, T. Onishi, and R. Kominami. 1979. *In vitro* synthesis of a 5S RNA precursor by isolated nuclei of rat liver and HeLa cells. *Cell*. 17: 163–173.
18. Hori, H., and S. Osawa. 1979. Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species. *Proc. Natl. Acad. Sci. U. S. A.* 76:381–385.
19. Jacq, B., R. Jourdan, and B. Jordan. 1977. Structure and processing of precursor 5S RNA in *Drosophila melanogaster*. J. Mol. Biol. 117:785–795.
20. Korn, L. J., E. H. Birkenmeier, and D. D. Brown. 1979. Transcription initiation of *Xenopus* 5S ribosomal RNA genes *in vitro*. *Nucl. Acids Res.* 7:947–958.
21. Korn, L. J., and D. D. Brown. 1978. Nucleotide sequence of *Xenopus borealis* oocyte 5S DNA: comparison of sequences that flank related eucaryotic genes. *Cell*. 15:1145–1156.
22. Kressmann, A., H. Hofstetter, E. DiCapua, and M. L. Birnstiel. 1979. A tRNA gene of *Xenopus laevis* contains at least two sites promoting transcription. *Nucl. Acids Res.* 7:1749–1763.
23. Levis, R. 1978. Kinetic analysis of 5S RNA formation in *Drosophila* cells. J. Mol. Biol. 122:279–283.
24. Maniatis, T., A. Jeffrey, and H. vande Sande. 1975. Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. *Biochemistry*. 14:3787–3794.
25. Martin, F. H., and I. Tinoco. 1980. DNA-RNA hybrid duplexes containing oligo (dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucl. Acids Res.* 8:2295–2299.
26. Ng, S.-Y., C. S. Parker, and R. G. Roeder. 1979. Transcription of cloned *Xenopus* 5S RNA genes by *X. laevis* RNA polymerase III in reconstituted systems. *Proc. Natl. Acad. Sci. U. S. A.* 76:136–140.
27. Ogden, R. C., J. S. Beckman, J. Abelson, H. S. Kang, D. Soll, and O. Schmidt. 1979. *In vitro* transcription and processing of a yeast tRNA gene containing an intervening sequence. *Cell*. 17:399–406.
28. Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* 13:319–353.
29. Rubin, G. M., and D. S. Hogness. 1975. Effect of heat shock on the synthesis of low molecular weight RNAs in *Drosophila*: accumulation of a novel form of 5S RNA. *Cell*. 6: 207–213.
30. Sakonju, S., D. F. Bogenhagen, and D. D. Brown. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5′ border of the region. *Cell*. 19:13–25.
31. Schmidt, O., J.-I. Mao, S. Silverman, B. Hovemann, and D. Soll. 1978. Specific transcription of eukaryotic tRNA genes in *Xenopus* germinal vesicle extracts. *Proc. Natl. Acad. Sci. U. S. A.* 75:4819–4823.
32. Silverman, S., O. Schmidt, D. Soll, and B. Hovemann. 1979. The nucleotide sequence of a cloned *Drosophila* arginine tRNA gene and its in vitro transcription in *Xenopus* germinal vesicle extracts. *J. Biol. Chem.* 254:10290–10294.
33. Sollner-Webb, B., and R. H. Reeder. 1979. The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in *X. laevis*. *Cell*. 18:485–499.
34. Sutcliffe, J. G. 1978. pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. *Nucl. Acids Res.* 5:2721–2728.
35. Sutcliffe, J. G. 1978. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harbor Symp. Quant. Biol.* 43:77–90.
36. Tanaka, T., B. Weisblum, M. Schnos, and R. B. Inman. 1975. Construction and characterization of a chimeric plasmid composed of DNA from *Escherichia coli* and *Drosophila melanogaster*. *Biochemistry*. 14:2064–2071.
37. Telford, J. L., A. Kressmann, R. A. Koski, R. Grosschedl, F. Muller, S. G. Clarkson, and M. L. Birnstiel. 1979. Delimitation of a promoter for RNA polymerase III by means of a functional test. *Proc. Natl. Acad. Sci. U. S. A.* 76:2590–2594.
38. Thimmappaya, B., N. Jones, and T. Shenk. 1979. A mutation which alters initiation of transcription by RNA polymerase III on the Ad5 chromosome. *Cell*. 18:947–954.
39. Wegnez, M., R. Monier, and H. Denis. 1972. Sequence heterogeneity of 5S RNA in *Xenopus laevis*. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 25:13–20.
40. Weil, P. A., J. Segall, B. Harris, S.-Y. Ng, and R. G. Roeder. 1979. Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. *J. Biol. Chem.* 254:6163–6173.
41. Weinmann, R., and R. G. Roeder. 1974. Role of DNA-dependent RNA polymerase III in the transcription of the tRNA and 5S RNA genes. *Proc. Natl. Acad. Sci. U. S. A.* 71:1790–1794.
42. Wensink, P. C., D. J. Finnegan, J. E. Donelson, and D. S. Hogness. 1974. A system for mapping DNA sequences in the chromosomes of *Drosophila melanogaster*. *Cell*. 3:315–325.
43. Wu, G. J. 1978. Adenovirus DNA directed transcription of 5.5S RNA in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 75:2175–2179.
44. Wu, G. J. 1980. Faithful transcription of adenovirus 5.5S RNA gene by RNA polymerase III in a human KB cell-free extract. *J. Biol. Chem.* 255:251–258.