Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3259–3263, June 1980
Biochemistry

Four sizes of transcript produced by a single sea urchin gene expressed in early embryos

(maternal messenger RNA/nuclear RNA/cloned gene)

AMY SHIU LEE*, TERRY L. THOMAS†, ZE'EV LEV†, ROY J. BRITTEN†‡, AND ERIC H. DAVIDSON†

*Department of Biochemistry, University of Southern California School of Medicine, Los Angeles, California 90033; and †Division of Biology, California Institute of Technology, Pasadena, California 91125

Contributed by Roy J. Britten, March 14, 1980

ABSTRACT **This report concerns a set of sea urchin egg and embryo transcripts complementary to a single-copy region of a cloned DNA fragment (Sp88). Three distinct 16-cell embryo polysomal RNA species were found to hybridize with this fragment. These RNAs are about 1700, 3000, and 4000 nucleotides (nt) in length, and the same species were identified in unfertilized eggs. A significant fraction of all three species of the egg and early embryo transcripts is polyadenylylated. At gastrula stage Sp88 transcripts are almost completely confined to the nucleus [Lev, Z., Thomas, T. L., Lee, A. S., Angerer, R. C., Britten, R. J. & Davidson, E. H. (1980) *Dev. Biol.* 75, in press]. The Sp88 transcripts of gastrulae are present as a fourth RNA species approximately 5800 nt in length. The four species share a sequence element of the cloned DNA fragment that is about 1000 nt long. These RNAs constitute a set of alternative partially overlapping transcripts from the same genomic region.**

The sea urchin egg contains a large store of maternal mRNAs, which are loaded on polysomes after fertilization and which account for most of the protein synthesis during early cleavage (reviewed in ref. 1). There are $\geq 10^4$ different species of maternal messages, each present in only one thousand to a few thousand copies per egg (2, 3), as well as a smaller number of more prevalent mRNAs (4). The mass of RNA in the embryo is more or less constant throughout early development (1), and a message present at 10^3 molecules per egg exists in about the same sequence concentration as does a typical rare polysomal mRNA at later stages. Thus, messages belonging to the rare sequence class are found at one or a few copies per cell at the 600-cell gastrula stage, and at the same or a slightly lower number of copies per cell in the 1500-cell pluteus stage embryo (2, 3).

Recently, we described a cloned single-copy DNA fragment that appears to contain part of a structural gene coding for a maternal mRNA of the rare sequence class (5). There are about 1400 RNA molecules per egg complementary to the single-copy DNA fragment, which is a 1.6-kilobase (kb) *Hae* III subfragment from a 5.6-kb genomic clone called Sp88. Five hours after fertilization, at the 16-cell cleavage stage, about 850 of these RNA molecules are located in the embryo cytoplasm, and at least half of these are polysomal. Later in development, however, the concentration of Sp88 RNA species in the polysomes (or cytoplasm) falls to <50 copies per embryo—i.e., less than 0.03 molecules per average cell at the pluteus stage. The Sp88 RNA is not detectable in adult intestine cytoplasmic RNA. Throughout development, this sequence nonetheless continues to be expressed in nuclear RNA (nRNA), where it is present at a level that is the same as the average level for the total rapidly turning over single-copy sequence transcripts in the nucleus at each stage. It is also represented in adult intestine nRNA. Thus, the Sp88 RNA belongs to the large class of rare embryonic messages whose presence in the cytoplasm appears to be regulated throughout developmental time (2,3) but whose sequences are found ubiquitously in nRNA (6). The present report concerns the nature of the transcripts complementary to the single-copy Sp88 *Hae* III (Sp88 H3) fragment in egg RNA, in 16-cell embryo polysomal RNA, and in gastrula total and nRNAs. We find that many of the Sp88 transcripts are polyadenylylated, and that there are at least three different but overlapping RNA species containing the Sp88 sequence in both egg and 16-cell polysomal RNAs. The gastrula nucleus contains a fourth species of Sp88 RNA, which is significantly larger than the three polysomal species observed.

MATERIALS AND METHODS

Preparation and Labeling of Plasmid DNA Probes. Sp88 is a 5.6-kb fragment of sea urchin DNA terminated by natural *Eco*RI sites, and inserted by means of these sites in plasmid RSF2124 (7, 8). Its origin, characteristics, and a partial restriction map were presented by Lev *et al.* (5). Many of the experiments described below were carried out with a 1.6-kb *Hae* III subfragment of Sp88, termed Sp88 H3 (see Fig. 1). Methods for the preparation, iodination, and strand separation of the Sp88 H3 fragment were also discussed earlier (5) and were applied in this work with the following minor modifications. After iodination at 60°C for 10 min, the reaction mixture was immediately passed over a Sephedex G-50 column equilibrated with 0.5 M NaCl/25 mM Hepes buffer (pH 6.5)/1 mM EDTA/0.1% $NaDodSO_4$. The excluded fraction was incubated for 2 hr at 60°C and precipitated in ethanol, and the strands were separated by electrophoresis in acrylamide/agarose gels. The separated DNA strands were electroeluted into dialysis tubing. The specific activity of the ^{125}I-labeled (^{125}I-) Sp88 H3 fragment prepared in this manner was $1–3 \times 10^7$ cpm/μg.

To facilitate isolation of the PstB, $PstA_1$, and $PstA_2$ subfragments (Fig. 1), the Sp88 H3 insert was subcloned. This was accomplished by blunt-end ligation of the Sp88 H3 fragment into pBR322 that had been cut with *Hin*cII to yield a pair of blunt ends at position 3907 in the sequence (9). There is a second *Hin*cII site in the pBR322 genome at position 652, and at this site the vector had previously been cut with *Sal* I, leaving single-stranded 4-nt overhanging ends that include this *Hin*cII site. This *Sal* site was reformed and sealed during ligation of the Sp88 H3 fragment. A restriction map of the subcloned Sp88 H3 insert and the immediately surrounding pBR322 sequence is shown in Fig. 1*B*. As indicated there, the PstB subfragment was isolated after digestion with *Hae* III and *Pst* I. The $PstA_1$ and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: nt, nucleotide(s); nRNA, nuclear RNA; kb, kilobase(s); HAP, hydroxyapatite; C_0t, moles of nucleotides per liter × sec; Me_2SO, dimethyl sulfoxide.

‡ Also Staff Member, Carnegie Institution of Washington.

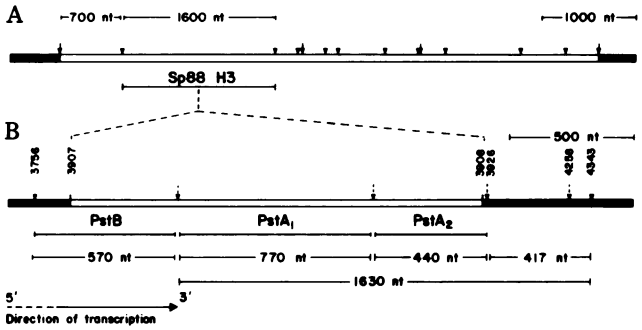

FIG. 1. Restriction endonuclease cleavage map of Sp88 and polarity of transcription. Restriction sites are designated as follows: ↓, *Eco*RI; ▼, *Hae* III; ↓, *Hha* I; ↓, *Bam*HI; and ↓, *Pst* I. The solid bar represents vector DNA. The fragments designated Sp88 H3, $PstA_1$, $PstA_2$, and PstB were used as probes to detect complementary RNAs in the experiments described. (*A*) Sp88 insert, bounded by natural *Eco*RI sites. Only *Eco*RI, *Hae* III, and *Bam*HI sites within the Sp88 insert are shown. (*B*) Restriction map of Sp88 H3 subclone and flanking plasmid sequences. *Hae* III, *Hha* I, and *Pst* I sites are shown. Numbers represent locations in the pBR322 nucleotide sequence according to Sutcliffe (9).

$PstA_2$ fragments were obtained by isolating the 1630-nt *Hae* III/*Pst* I fragment of the Sp88 H3 subclone (see Fig. 1) and then redigesting this with *Hha* I. The two Sp88 fragments released by this enzyme ($PstA_1$ and $PstA_2$) were separated from each other and from the remaining 332-nt plasmid sequence by gel electrophoresis and electroeluted as described above. For use in the experiments described, these fragments were nick translated (10) to a specific activity of $2–4 \times 10^8$ cpm/μg.

Preparation of RNA. Total RNA was isolated from sea urchin eggs, 16-cell embryos, and gastrula embryos as described (3). Sixteen-cell embryo polysomal RNA was prepared from 5-hr embryos demembranized at fertilization with papain (11). The embryos were lysed by suspension in Ca^{2+}- and Mg^{2+}-free seawater (pH 7.7), and the lysate was loaded on sucrose gradients, as described by Galau *et al.* (2). The polysomal fractions were collected, urea was added to a final concentration of 7 M, and the RNA was immediately extracted. Polyadenylylated RNA fractions were obtained by standard oligo(dT) column chromatography. The RNA was bound in 10 mM Tris (pH 7.4)/1 mM EDTA/0.5% $NaDodSO_4$/0.5 M NaCl and eluted in the same buffer without NaCl (12). About 2% of the mass of total cellular RNA was bound. Gastrula nRNA was prepared by published procedures (13). Briefly, the embryos were sheared in a Waring Blendor, the nuclei were pelleted and lysed in urea buffer, and the RNA was extracted by standard phenol procedures.

RNA·DNA Hybridization and Assays. Procedures used in the titration of the iodinated upper-strand fragment of Sp88 H3 with various RNAs were described in detail by Lev *et al.* (5). After hybridization, the samples were diluted with 1 ml of 0.12 M phosphate buffer and 0.05% $NaDodSO_4$, and were passed over 2-ml hydroxyapatite (HAP) columns operated at 50°C in 0.12 M phosphate buffer/0.05% $NaDodSO_4$. The bound hybrids were eluted by washing the column with three 3-ml portions of 0.12 M phosphate buffer/0.05% $NaDodSO_4$ at 99°C. RNA·DNA hybridizations of the nick-translated tracers with egg and gastrula RNAs were also assayed by binding to HAP (2, 14). After incubation, the total duplex bound to HAP was measured, and the DNA·DNA duplex due to DNA tracer self-reaction was measured by HAP binding after digestion with ribonuclease under low salt conditions (2). Where hybrid formation was assayed by S1 nuclease resistance rather than HAP binding, the hybridization reactions were carried out in 0.75 M NaCl. After incubation, these reaction mixtures (volume ≈12 μl) were diluted into 2 ml of 0.15 M NaCl/0.03 M sodium acetate (pH 4.5)/1 mM $ZnSO_4$/sheared calf thymus DNA at 5 μg/ml. One half of the sample was used to monitor the total amount of trichloroacetic acid-precipitable tracer present before digestion, and the other half was digested with 2 μl of S1 nuclease (Calbiochem, B grade), which was sufficient to degrade the single-stranded DNA but leave the double-stranded DNA intact. After 1 hr at 37°C, trichloroacetic acid was added to 10% and the precipitate was collected on GF/C filters and its radioactivity was measured.

RESULTS

Direction of Transcription. Only one strand of the expressed single-copy region of the Sp88 insert reacts with either sea urchin egg RNA or with cytoplasmic, polysomal, or nuclear embryo RNAs. This was established earlier (5), by separating the strands of the Sp88 H3 fragment and testing each strand individually with the various RNA preparations. The reactive strand was, in all cases, the more slowly migrating or "upper" strand under neutral gel electrophoresis conditions.

The following experiment established the direction of transcription with respect to the restriction map of Sp88 shown in Fig. 1*A*. The 5.6-kb *Eco*RI insert of Sp88 was isolated and the 5′ ends were labeled with polynucleotide kinase. The fragment was next cleaved with *Bam*HI. As shown in Fig. 1, the longest of the two labeled subfragments resulting from the *Bam*HI digestion (2.7 kb) contained the left-hand terminus of the original Sp88 insert. This fragment includes the transcribed (1.6-kb) Sp88 H3 subfragment. The approach followed was to determine whether the 3′ end of the upper strand of the Sp88 H3 probe fragment—i.e., the transcribed strand—lies toward the left or the right of the insert as it is oriented in Fig. 1. Driver quantities of the separated Sp88 H3 strands were prepared and were allowed to react in large excess with the asymmetrically labeled, 2.7-kb *Bam*HI tracer, after denaturation of the latter. Only the upper strand reacted with this tracer (39% upper strand bound to HAP at driver C_0t of 1.2×10^{-3} moles of nucleotide per liter × sec vs. no reaction at all with the lower strand). It follows that the 3′ end of the transcribed strand is to the left in Fig. 1, and, therefore, transcription must proceed in a left-to-right direction, as shown. Important additional evidence is that under the usual conditions, no restriction fragments downstream (i.e., to the right in Fig. 1) from the Sp88 H3 fragment react significantly with egg RNA, whereas the 0.7-kb *Hae* III fragment upstream at the left-hand terminus of the insert is clearly represented in egg RNA (data not shown).

Some Sp88 Transcripts Are Polyadenylylated. About half of the maternal mRNA in the sea urchin egg is polyadenylylated by the criterion of oligo(dT)-cellulose binding (reviewed in ref. 1), and this is true of both the rare and the more prevalent classes of stored message. Poly(A)-containing RNA from unfertilized eggs includes essentially all the sequence diversity of unfractionated egg RNA or of early embryo polysomal RNA.§ The experiments summarized in Table 1 show that a significant portion of the Sp88 transcripts of egg and 16-cell embryos is also polyadenylylated. $Poly(A)^+$ and $poly(A)^-$ RNA fractions were prepared by the standard oligo(dT) affinity column procedures and the relative concentrations of Sp88 transcripts in the $poly(A)^+$ and $poly(A)^-$ RNAs were measured by titration (5). The probe was ^{125}I-Sp88 H3 upper strand. We found an approximately 30-fold enrichment of Sp88 transcripts in the

§ F. D. Costantini, R. J. Britten, and E. H. Davidson, unpublished data.

Table 1. Number of Sp88 RNA molecules included in poly(A)$^+$ and poly(A)$^-$ RNA fractions per egg or 16-cell embryo

RNA	poly(A)$^+$ RNA	poly(A)$^-$ RNA	Minimum fraction polyadenylylated
Unfertilized egg	510	890	0.36
16-cell embryo	330	510	0.39

Total RNA was extracted from eggs or 16-cell (5-hr) embryos and fractionated by oligo(dT) column chromatography. Upper strand Sp88 H3 DNA was labeled by iodination and used to titrate the poly(A)$^+$ RNA and poly(A)$^-$ RNA fractions, as described by Lev *et al.* (5) and in *Materials and Methods*. These experiments showed that there was a 28-fold enrichment of Sp88 sequences in poly(A)$^+$ egg RNA and a 31-fold enrichment of these sequences in poly(A)$^+$ 16-cell RNA, relative to the respective poly(A)$^-$ RNA. Let: C be the ratio of Sp88 transcript concentrations in poly(A)$^+$/poly(A)$^-$ RNA, i.e., 28 and 31, respectively; f be the fraction in total RNA of the poly(A)$^+$ preparations used for these titrations; T be the total number of Sp88 transcripts per egg or 16-cell embryo; and A be the number of polyadenylylated transcripts per egg or 16-cell embryo. For all polyadenylylated molecules included in the poly(A)$^+$ fractions employed in these experiments, $C = (A/f)/[(T - A)/(1 - f)]$; thus, $A = T/[(1 - f)/(f \cdot C) + 1]$. For the egg, T = 1400 molecules, and for the 16-cell embryo, T = 840 molecules; f was 0.02 for both RNA preparations. The fractions of transcripts thus calculated as polyadenylylated are listed as minimum, because poly(A) tails could have been lost from some molecules by breakage or degradation.

poly(A)$^+$ RNA fractions of both unfertilized eggs and 16-cell embryos, relative to the poly(A)$^-$ RNA fractions. The number of Sp88 transcripts per egg or 16-cell embryo included in the poly(A)$^+$ RNA fraction was then calculated (see legend to Table 1). The results show that at least 35–40% of the Sp88 transcripts are polyadenylylated. These values are clearly minimum ones, because the RNAs could have been partially degraded during extraction and purification or the poly(A) tails of some molecules could have been too short to permit binding to oligo(dT) columns under the conditions used (4, 15). While the RNA used for these experiments displayed optical absorbance profiles of undegraded RNA, it cannot be excluded that the Sp88 transcripts were preferentially sensitive. However, it is also possible that the Sp88 transcript population is naturally heterogeneous with respect to the possession of long poly(A) tails. Costantini *et al.*§ showed that, on the average, 75% of the molecules of about half the species of egg RNA are polyadenylylated by the criterion of oligo(dT)-cellulose binding. The rest of the maternal mRNA sequence set is found also entirely on nonpolyadenylylated molecules. Thus, Table 1 indicates that the Sp88 transcripts are polyadenylylated, by the same criterion, to an extent that is not unusual for the rare class of sea urchin egg maternal mRNAs.

Multiple Sp88 Transcript Species. Sp88 transcripts have been found in at least three different biological contexts (5): as stored maternal mRNAs, as polysomal early embryo messages presumably being translated, and as nRNAs in cells that do not significantly export these mRNAs to the cytoplasm. To compare the Sp88 transcripts in these three conditions, we extracted total RNA from unfertilized egg and gastrula stage embryos and polysomal RNA from 16-cell embryos. Maximum care was taken to prevent degradation. These RNAs were resolved by velocity sedimentation through denaturing [80% (vol/vol) dimethyl sulfoxide (Me_2SO)] sucrose gradients. Fractions collected from each gradient were individually allowed to react with excess ^{125}I-SP88 H3 upper strand probe. The amount of tracer hybridized by the RNA in each fraction indicates the amount of Sp88 transcripts in that RNA size class. The size distribution of Sp88 transcripts measured in this way is shown in Fig. 2. Surprisingly, there appear to be three distinct classes of transcript reacting with the single-copy Sp88 H3 probe in both unfertilized egg RNA and 16-cell embryo polysomal RNA. These transcripts are about 1700, 3000, and 4000 nt long in both RNAs, as determined with reference to the optical absorbance profile of the ribosomal RNAs in the same gradients. In order of increasing length, the molar ratio of the three transcript

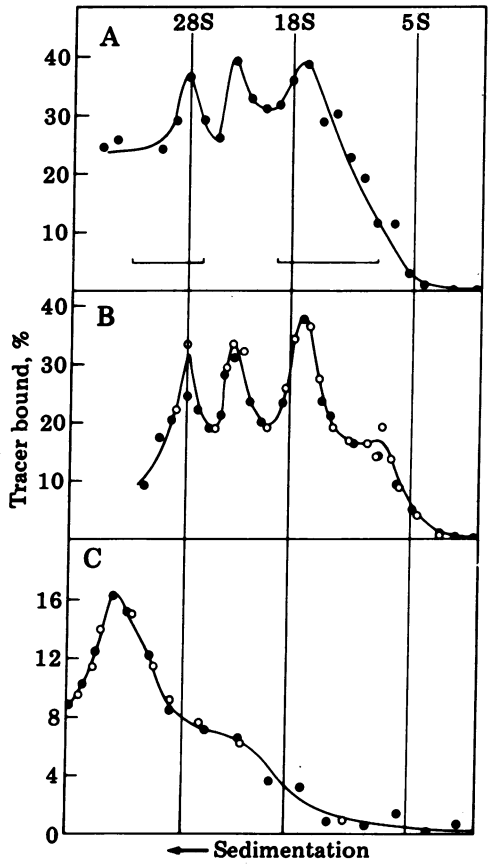

FIG. 2. Size of RNA molecules hybridizing with Sp88 H3 upper strand probe. Sea urchin RNA was fractionated by sedimentation in denaturing Me_2SO gradients (16). The conditions used are completely denaturing for RNA molecules (17). RNA (300 μg) was dissolved at 1.8 mg/ml in 80% (vol/vol) Me_2SO/0.1 M LiCl/0.01 M Tris-HCl (pH 6.5)/0.005 M EDTA/0.2% $NaDodSO_4$ and heated at 55°C for 5 min. This RNA solution was then diluted with the same buffer to 1.2 mg/ml and 60% Me_2SO, and was sedimented at 28°C, 35,000 rpm through a 4–20% exponential sucrose gradient containing 60% Me_2SO for 65 hr in a Beckman SW 41 Ti rotor. The gradient was pumped through an ISCO recording spectrophotometer and about 30 fractions (≈0.35 ml per fraction) were collected. The three major absorbance peaks (at 260 nm) contained 5S, 18S, and 28S ribosomal RNAs, and these peaks served as internal size markers in each of the gradients. Sizes of the sea urchin ribosomal RNAs and the Sp88 transcript peaks were calculated with reference to *Escherichia coli* ribosomal RNA markers, the length of which is exactly known (18, 19). These were run together with sea urchin ribosomal RNA in a denaturing (methylmercury) gel in an independent experiment. The ribosomal RNAs are: 18S, 1840 nt; 28S, 4000 nt. The RNA fractions were adjusted to 0.3 M sodium acetate (pH 6.8), and 20 μg of yeast RNA was added to each fraction as carrier for precipitation in 66% (vol/vol) ethanol at −20°C. After centrifugation, each RNA fraction was allowed to react with excess Sp88 H3 upper strand, as follows: Each RNA precipitate was resuspended in 12 μl of 0.5 M phosphate buffer containing 1200 cpm of ^{125}I-Sp88 H3 DNA (specific activity 3×10^7 cpm/μg) and sealed in a glass capillary. These RNA fractions were heated to 105°C for 2 min and incubated at 60°C to DNA C_0t 6×10^{-3} mol·sec/liter. The amount of ^{125}I-DNA in hybrid was assayed by binding to HAP in 0.12 M phosphate buffer at 50°C. (*A*) Total mature oocyte RNA, (*B*) 16-cell polysomal RNA, and (*C*) total gastrula RNA. ○ and ● in *B* and *C* represent separate experiments carried out with independent RNA preparations.

species is about 2:1:1 in the 16-cell embryo polysomal RNA. Though slightly less well resolved, the three Sp88 transcript species appear to be present in the unfertilized egg RNA in about the same molar ratios. It is not likely that these three species are an artifact of degradation, considering the apparently intact condition of the RNA preparations layered on the gradients, the discrete character of the transcript peaks, and their excellent reproducibility with separate, unrelated RNA preparations, as shown in Fig. 2*B*. Poly(A)$^+$ RNA with a weight-average length of 5000 nt can be extracted from unfertilized eggs.§ Furthermore, the same multiple transcript species can be observed in RNA gel blots (20). Experiments of this kind (unpublished) confirm the existence of the three RNA species shown in Fig. 2 and demonstrate that all three species are polyadenylylated. In addition, a fourth and longer transcript may exist in the egg poly(A)$^+$ RNA, but it would not have been resolved on the gradient shown in Fig. 2. Because the same three species that are resolved in egg RNA are also in the polysomal RNA at the 16-cell embryo stage, there are no *a priori* grounds for considering the two larger species to be precursors to the 1700-nt species or incomplete processing products. Nor can nonpolysomal cytoplasmic contaminants easily account for the diversity of Sp88 RNA transcripts, because at least half of these transcripts are polysomal in the 16-cell embryo as opposed to none in the unfertilized egg, while the ratio of transcript sizes is about the same in the two RNA preparations. The same evidence argues against the need for major changes in the transcript size during mobilization and assembly of the maternal SP88 transcripts into embryo polysomes, though of course many kinds of minor alterations could not have been detected.

By gastrula stage, the Sp88 H3 sequence is no longer significantly represented in polysomal RNA, though it continues to be expressed in the nucleus. The Sp88 transcripts in total gastrula RNA are at least 90% nuclear (5). Fig. 2*C* shows that in the gastrula most of these transcripts are about 5800 nt long and some could be even larger, beyond the resolving range of the gradient. Some of the shorter transcript forms seen earlier could also be present. However, it is clear that the major Sp88 RNA in the gastrula nucleus is a significantly longer species, which includes a region of homology to some portion of the same single-copy DNA probe as the three mRNA species identified above.

Sequence Overlap of the Sp88 Transcript Species. Two experiments were carried out to determine whether the various species of Sp88 transcript include different regions of the probe sequence. In the first, shown in Table 2, total egg RNA, isolated "16S" and "28S" egg RNA fractions, and total gastrula RNA were allowed to react with separated strands of ^{125}I-SP88 H3 tracer. When the fraction of the tracer sequence protected from S1 nuclease digestion was measured, the amount of the probe hybridized was found to be approximately the same for all the RNAs tested. Lev *et al.* (5) reported earlier that about 1100 nt of the Sp88 H3 sequence are resistant to S1 nuclease digestion after hybridization with egg RNA. Table 2 shows that the individual "16S" and "28S" RNA fractions, containing mainly the 1700- and 4000-nt species, respectively, protected similar lengths of sequence, about 1000 nt. Gastrula RNA protected slightly less of the tracer in this experiment but the difference is probably not significant. These results suggest but do not prove that the various Sp88 transcript species all include approximately the same element(s) of the Sp88 H3 sequence. Because transcription proceeds into this fragment from the left in the orientation shown in Fig. 1, and because the adjacent 700-nt *Hae* III fragment to the left is also represented in egg RNA whereas sequences downstream are not, the simplest interpretation is that the various forms of Sp88 transcript share

Table 2. Protection of Sp88 H3 sequence from S1 nuclease digestion by hybridization with various RNAs

Hybridization with	^{125}I-Sp88 H3 upper strand resisting S1 nuclease, %
Total egg RNA	53
"28S" egg RNA	55
"16S" egg RNA	54
Total gastrula RNA	48
Sp88 H3 DNA	85

The "28S" and "16S" RNA fractions are indicated by the brackets in Fig. 2*A*. The Sp88 RNA to Sp88 H3 DNA molar sequence ratios in these reactions, based on titration measurements, were 1.5–1.8, and were sufficient so that higher ratios yielded no further reaction (5). The reactions were run to RNA C_0t 1.4×10^4 mol·sec/liter. Because the Sp88 H3 fragment is 1600 nt in length, and 85% of ^{125}I in the reactions could be covered as S1 nuclease-resistant duplex after reaction with homologous DNA, the length of sequence represented by the RNAs is, according to this experiment, $(0.55 \times 1600)/0.85$ or about 1000 nt.

the left-hand 1000 nt of the Sp88 H3 fragment. Thus, the right end of this fragment should not be represented in any of the RNAs. To test this expectation, the experiment summarized in Table 3 was carried out. Three restriction enzyme subfragments consisting of the left (PstB), the center ($PstA_1$), and the right ($PstA_2$) portions of the Sp88 H3 sequence were isolated. These fragments are identified in Fig. 1, and their preparation from subcloned Sp88 H3 DNA is summarized in *Materials and Methods*. The fragments were nick translated to high specific activity and were allowed to react separately with egg RNA and with gastrula nRNA. High concentration of RNA, rather than strand separation, was relied on to suppress tracer self-reaction in this experiment, and hybrid formation was assayed by HAP binding. The tracer reactivities were measured by reaction with excess Sp88 DNA, which for short randomly terminated tracers may significantly overestimate the reactivity with RNA as measured by HAP binding (2). For this reason, the experiment of Table 3 is of limited quantitative accuracy, but the qualitative result is unequivocal. The left-hand or PstB fragment is represented more or less completely in both the nuclear Sp88

Table 3. Hybridization of egg RNA and gastrula nRNA with Sp88 H3 subfragments

Sp88 subfragments*	Length of sea urchin DNA sequence, nt	Hybridization, %		
		Max. possible†	Observed (% of max.‡)	
			Egg RNA	Gastrula nRNA
PstB	420	32	32 (100)	28 (88)
$PstA_1$	770	38	16 (42)	14 (37)
$PstA_2$	440	24	1	1

* See Fig. 1*B*. Fragments were prepared as described in *Materials and Methods*.

† Reactivities of the three tracer preparations with excess DNA were, respectively, 88%, 77%, and 79%. Their total lengths (including plasmid sequence) were (Fig. 1) 570 nt, 770 nt, and 440 nt. To calculate the maximum hybridization, it is assumed that (as shown) only one DNA strand is represented in RNA, and that the tracer fragment length is very short, so that the results are comparable to an S1 nuclease assay. Tracers nick-translated to the specific activity used here ($2–4 \times 10^8$ cpm/μg) are 100–200 nt long. In the case of the $PstA_2$ fragment, a 332-nt plasmid sequence (Fig. 1*B*) contaminated the preparation, so that only 48% of the tracer was sea urchin DNA.

‡ Hybridization conditions are given in *Materials and Methods*. The calculated Sp88 RNA·DNA molar sequence ratios for the PstB fragment were 10:1 for both the egg RNA and the gastrula nRNA. RNA·DNA molar sequence ratios were the same for all fragments. Reactions were run to RNA C_0t 13,000 mol·sec/liter.

gastrula transcripts and in the egg Sp88 transcripts, the right-hand or $PstA_2$ fragment is not at all represented in either of these RNAs, and the middle or $PstA_1$ fragment is partially represented in both RNAs.

We conclude that the cytoplasmic and nuclear SP88 transcripts overlap in that all four species contain approximately 1000 nt of sequence at the left-hand end of the Sp88 H3 fragment. The site or sites where homology with RNA ends (in the $PstA_1$ fragment) may well represent the 3′ end of all four transcript species [exluding the poly(A) sequences]. Alternatively, this region could include the 5′ terminus of an intervening sequence whose 3′ terminus lies far downstream, beyond the limits of the whole SP88 insert. This alternative seems less likely, because it requires that the 5800-nt gastrula nuclear species is a processed form, even though there is no significant cytoplasmic export of Sp88 transcripts at this stage. Another possible explanation for the multiple overlapping transcripts is that these derive from several copies of the Sp88 H3 sequence in the genome. We regard this as extremely unlikely in view of extensive evidence from renaturation kinetics, genome blots, and λ library screens that this sequence occurs only once per haploid genome (data to be presented elsewhere). Therefore, the differences in length between the four Sp88 species probably is due to the existence of several alternative processing pathways, or to alternative transcriptional initiation sites.

CONCLUSIONS

Almost all animal cells that have been examined display a prominent rare message class (21), but due to the difficulty of studying transcripts present in low sequence concentration relatively little is known about these mRNAs. An opportunity to investigate an apparently typical example of a rare sea urchin embryo message is provided by the cloned Sp88 single-copy probe (5). We were surprised to discover that there are at least three different poly(A) species, about 1700, 3000, and 4000 nt in length, which are probably polysomal in the 16-cell embryo and which share the same Sp88 H3 sequence element. This situation is reminiscent of the alternative processing pathways that are known to produce diverse but partially overlapping mRNAs in several animal virus systems (e.g., refs. 22–26), as well as in a murine cell line (27). Whether or not there is a similar cause for the appearance of the several Sp88 transcription species, these examples suggest that multiple utilization of the same codogenic sequence is not to be regarded primarily as a device to make "efficient" use of limited genetic information, a teleological argument sometimes applied to the viral examples. An enormous complexity, about 3.7×10^7 nt of single-copy sequence, has been measured for sea urchin maternal message (2, 3). The case of Sp88 raises the possibility that the actual diversity of the proteins coded by these messages could be significantly higher than has been calculated on the simple assumption that each distinct RNA sequence codes for a single protein species.

Sp88 sequences essentially disappear from the polysomes after the earliest stages of development, but the sequence continues to be represented in nRNA. The Sp88 nuclear transcripts are always present at roughly the same level as is the average single-copy sequence; at gastrula stage this is close to one copy per nucleus. Preliminary unpublished observations indicate that, like most heterogeneous nRNA, the gastrula Sp88 nuclear transcripts are synthesized and rapidly turned over. Tables 2 and 3 of this paper show that the same 3′ Sp88 H3 sequence element shared by the three message species is present in the large nuclear species. One possibility is that ubiquitous nuclear transcripts that include structural gene sequences actually have nothing to do with mRNA production, and the latter is indeed transcriptionally regulated. In this case, we might expect that the initiation site(s) for the "nonproductive" 5800-nt gastrula Sp88 nuclear transcript would be in a different location from that utilized during oogenesis, when Sp88 transcripts destined for polysomal expression are being synthesized. On the other hand, if Sp88 transcription always begins at the same place and occurs at the same rate, then stage-specific differences in posttranscriptional processing will be required to explain the greater than 10-fold changes observed in the cytoplasmic expression of this sequence during development.

This research was supported by National Institutes of Health Grant HD-05753. A.S.L. was supported by an American Cancer Society, California Division, Senior Fellowship (D-353). T.L.T. was supported by an ACS, California Division, Lievre Fellowship (J-409). Z.L. was supported by a Chaim Weizmann Fellowship and by a Gosney Fellowship.

1. Davidson, E. H. (1976) *Gene Activity in Early Development* (Academic, New York).
2. Galau, G. A., Klein, W. H., Davis, M. M., Wold, B. J., Britten, R. J. & Davidson, E. H. (1976) *Cell* **7**, 487–505.
3. Hough-Evans, B. R., Wold, B. J., Ernst, S. G., Britten, R. J. & Davidson, E. H. (1977) *Dev. Biol.* **60**, 258–277.
4. Wilt, F. H. (1977) *Cell* **11**, 673–681.
5. Lev, Z., Thomas, T. L., Lee, A. S., Angerer, R. C., Britten, R. J. & Davidson, E. H. (1980) *Dev. Biol.* **75**, in press.
6. Wold, B. J., Klein, W. H., Hough-Evans, B. R., Britten, R. J. & Davidson, E. H. (1978) *Cell* **14**, 941–950.
7. So, M., Gill, R. & Falkow, S. (1975) *Mol. Gen. Genet.* **142**, 239–249.
8. Lee, A. S., Britten, R. J. & Davidson, E. H. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42**, 1065–1076.
9. Sutcliffe, J. B. (1978) *Nucleic Acids Res.* **5**, 2721–2728.
10. Maniatis, T., Jeffrey, A. & Van de Sande, H. (1975) *Biochemistry* **14**, 3787–3794.
11. Hynes, R. O. & Gross, P. R. (1972) *Biochim. Biophys. Acta* **259**, 104–111.
12. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
13. Smith, M. J., Hough, B. R., Chamberlin, M. E. & Davidson, E. H. (1974) *J. Mol. Biol.* **85**, 103–126.
14. Galau, G. A., Britten, R. J. & Davidson, E. H. (1974) *Cell* **2**, 9–21.
15. Dubroff, L. M. & Nemer, M. (1976) *Nature (London)* **260**, 120–124.
16. Costantini, F. D., Scheller, R. H., Britten, R. J. & Davidson, E. H. (1978) *Cell* **15**, 173–187.
17. Strauss, J. H., Kelly, R. B. & Sinsheimer, R. L. (1968) *Biopolymers* **6**, 793–807.
18. Brosius, J., Palmer, M. L., Kennedy, P. J. & Noller, H. F. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4801–4805.
19. Brosius, J., Dull, T. J. & Noller, H. F. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 201–204.
20. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350–5354.
21. Davidson, E. H. & Britten, R. J. (1979) *Science* **204**, 1052–1059.
22. McGrogan, M. & Raskas, H. J. (1977) *J. Virol.* **23**, 240–249.
23. Nevins, J. R. & Darnell, J. E. (1978) *J. Virol.* **25**, 811–823.
24. Ziff, E. & Fraser, N. (1978) *J. Virol.* **25**, 897–906.
25. Weinberg, R. A., Ben-Ishai, Z. & Newbold, J. E. (1974) *J. Virol.* **13**, 1263–1273.
26. Kamen, R. & Shure, H. (1976) *Cell* **7**, 361–371.
27. Perry, R. P. & Kelley, D. E. (1979) *Cell* **18**, 1333–1339.