

## Developmental regulation of a gastrula-specific gene injected into fertilized *Xenopus* eggs

P.A.Krieg and D.A.Melton

Department of Biochemistry and Molecular Biology, Harvard University,  
7 Divinity Avenue, Cambridge, MA 02138, USA

Communicated by R.Cortese

**To study the transcriptional regulation of genes during early *Xenopus* development, we have isolated a gene that is first transcribed at the mid-blastula transition. Transcription of this gene, called GS17, stops at mid-gastrula and the mRNA is rapidly degraded. Consequently, transcripts of GS17 are only present for a brief period, primarily during gastrulation. When the GS17 gene is injected into fertilized eggs, transcription from the injected DNA mimics the expression pattern of the endogenous gene, i.e., both the switch-on and switch-off of transcription are correctly regulated. The injected DNA is not significantly amplified and remains extrachromosomal. The correct expression of genes injected into *Xenopus* eggs will make it possible to investigate maternal factors involved in activating the embryonic genome.**

**Key words:** development/gastrula/egg microinjection/*Xenopus*

### Introduction

After fertilization, the frog egg enters a period of rapid cleavage giving rise to a blastula of 4000–8000 cells. This early development is directed entirely by the maternal genome; all the required proteins are provided either from a maternal stock or by translation of stored maternal mRNAs. When the embryo reaches the mid-blastula stage, changes occur in cell cycle timing, cell motility and the zygotic genes are transcribed for the first time (Signoret and Lefresne, 1971; Brown and Littna, 1964; Newport and Kirschner, 1982a). These changes, called the mid-blastula transition or MBT (Signoret and Lefresne, 1971; Gerhart, 1980; Newport and Kirschner, 1982a) commence ~7 h after fertilization. Thus, in the frog embryo, studies on how embryonic genes are selectively activated should be facilitated by the fact that it is possible to obtain large numbers of embryos before and after genes are first transcribed. In effect, eggs and cleavage stage embryos comprise a pure population of precursor cells and blastula stage are their daughters that have selectively turned on some genes.

To analyze the transcriptional activation of embryonic genes, we reasoned that it would be helpful to isolate a gene that is turned on at the MBT and whose transcripts are absent from oocytes and eggs. Previous studies have shown that the vast majority of genes transcribed at the MBT are also expressed during oogenesis, i.e., their transcripts are found in the maternal RNA pool (Dworkin and Dawid, 1980; Colot and Rosbash, 1982). Nonetheless, differential screening of a *Xenopus* gastrula cDNA library has enabled us and others (Sargent and Dawid, 1983) to select clones that are not found in the maternal RNA pool, but are expressed very early when the embryonic genome is activated. Of the sequences identified by this method, we have chosen one, called GS17, to study in detail for two reasons. First, GS17 is

among the first protein-coding genes to be expressed during development and therefore can be used to study the transcriptional activation of genes at the MBT. Secondly, GS17 has the unusual property that its transcripts are found only at a critical time, principally during gastrulation. This gastrula-specific expression pattern leads to the suggestion that the GS17 gene product may play a role in directing gastrulation or determining cell fates.

We would like to know how genes like GS17 are transcriptionally activated at the MBT whereas most other genes are turned on at a later stage. In other animals, notably flies (Spradling and Rubin, 1983; Goldberg *et al.*, 1983), sea urchins (Davidson, personal communication) and mice (Palmiter and Brinster, 1985; Krumlauf *et al.*, 1985), the *cis*-acting signals involved in this type of transcriptional control have been very effectively studied by re-injecting altered genes into developing embryos. To date this approach has had very limited success in frogs and in no case has the correct developmental expression of a RNA polymerase II gene been demonstrated. It is worth noting that most previous studies with developing frog embryos used heterologous genes, from rabbits (Rusconi and Schaffner, 1981), sea urchins (Bendig, 1981; Etkin *et al.*, 1984), *Drosophila* and animal viruses (Etkin *et al.*, 1984; Etkin and Balcells, 1985), and these genes may, therefore, have been inappropriate tests. In some experiments homologous *Xenopus* adult and larval globin genes were injected (Bendig and Williams, 1983, 1984), but correct developmental regulation was not observed perhaps because these genes are normally expressed in a tissue-specific manner relatively late in embryonic life.

The most positive results using the *Xenopus* embryo as an expression system have been obtained by Busby and Reeder (1983) who showed that genes transcribed by RNA polymerase I, the *Xenopus* rRNA genes, are turned on at the same time (late blastula) as the endogenous ribosomal genes when these genes are injected into fertilized eggs. Their results suggest that *Xenopus* genes which are normally expressed early in embryogenesis may also be correctly regulated when they are re-introduced into the developing embryo. With this in mind we have injected cloned copies of GS17 into fertilized eggs to determine whether RNA polymerase II genes are correctly regulated during subsequent embryonic development. Results presented here show that both the switch-on and switch-off of transcription of injected GS17 sequences is accurately controlled in the developing embryo. This correct developmental expression does not require special P-element or retrovirus vectors and occurs using standard plasmid vectors that remain extrachromosomal.

### Results

#### *Isolation of sequences expressed early in development*

Though transcription of the embryonic genome is first observed at the mid-blastula stage (Brown and Littna, 1964; Newport and Kirschner, 1982a), we prepared a cDNA library from embryos at a slightly later stage of development so that new transcripts

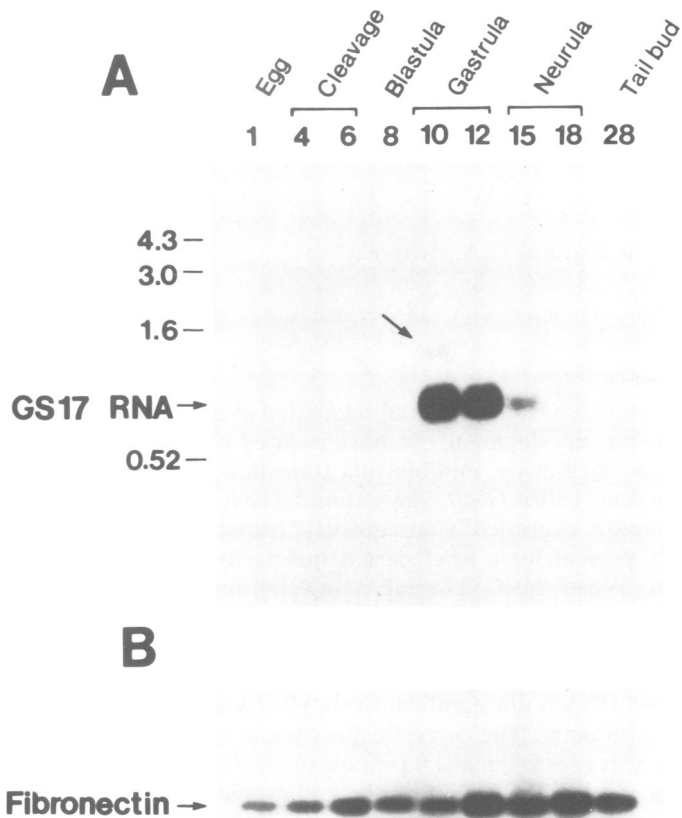
could accumulate to detectable levels. Using poly(A)<sup>+</sup> RNA isolated from mid-gastrula (stage 11) embryos we constructed a cDNA library in the bacteriophage vector  $\lambda$ gt10 that contains  $\sim 1 \times 10^6$  recombinants with average insert size of 1.1 kb. Clones corresponding to mRNAs transcribed starting at the MBT were selected from this gastrula library by differential screening. Since the egg contains a large amount of mRNA and most of these maternal sequences are stable throughout early development, only a small percentage of the recombinants in the gastrula library are derived from newly transcribed mRNAs (Colot and Rosbash, 1982; Dworkin and Dawid, 1980; Sargent and Dawid, 1983). To detect this small number of new sequences, replica filters containing a total of 20 000 recombinant bacteriophage were probed with [<sup>32</sup>P]cDNA prepared either from defolliculated oocyte poly(A)<sup>+</sup> RNA or from gastrula stage embryo poly(A)<sup>+</sup> RNA. Those clones that gave a weak or undetectable response with oocyte cDNA but were strongly positive for gastrula cDNA were purified and subjected to further characterization. Of the 101 recombinants selected, 53 were found to contain mitochondrial sequences by hybridization with probe prepared from *Xenopus* mitochondrial DNA (Rastl and Dawid, 1979). These clones were probably identified because the mitochondrial genome is transcriptionally active during the cleavage stages of development (Chase and Dawid, 1972; Young and Zimmerman, 1973).

The 48 remaining cDNA sequences were further characterized in RNA blotting experiments. cDNA inserts from each of the  $\lambda$ gt10 recombinants were excised, labelled by nick translation and used as probe against either oocyte or gastrula poly(A)<sup>+</sup> RNA in a Northern blot. Under these sensitive screening conditions most of the sequences that were apparently absent from oocytes by the plaque-screening criterion, could be detected in poly(A)<sup>+</sup> RNA oocyte RNA, although generally at a very low level (data not shown). Most of the 48 cDNA clones correspond to RNAs that are 10- to 20-fold more abundant in gastrula compared with oocytes. We detected only a single cDNA sequence, GS17, that is present in gastrula RNA and completely absent from oocyte RNA.

As noted in previous experiments where libraries of *Xenopus* oocyte or embryo mRNA sequences have been screened, only a proportion of the total number of recombinants examined are detectable when probed with homologous cDNA (Golden *et al.*, 1980; Dworkin and Dawid, 1980). In our case, only 35–40% of the clones (a total of 7000–8000 plaques), were detected after a 3-day autoradiographic exposure, indicating that the majority of sequences present in the library represent very rare mRNAs. Therefore, although GS17 is the only sequence we isolated that is present in gastrula but absent from the oocyte, we are aware that many other mRNAs may be similarly expressed but would not be detected by our screening procedure.

#### Expression of GS17 during embryonic development

To examine the expression of the GS17 gene during embryogenesis, Northern blots were performed with RNA isolated from a series of developmental stages ranging from the fertilized egg to the tailbud stage (Figure 1). The GS17 probe hybridizes to an RNA transcript  $\sim 750$  bases in length. While no GS17 mRNA is observed at any developmental stage up to and including stage 8, it is already abundant by stage 10, early in gastrulation. The amount of GS17 RNA remains constant until late in gastrulation (stage 12), and then diminishes rapidly until none is visible at any developmental stage after neurulation has occurred. The stage 28 tailbud embryo contains many fully differentiated cell types including blood, muscle, nerve and skin,

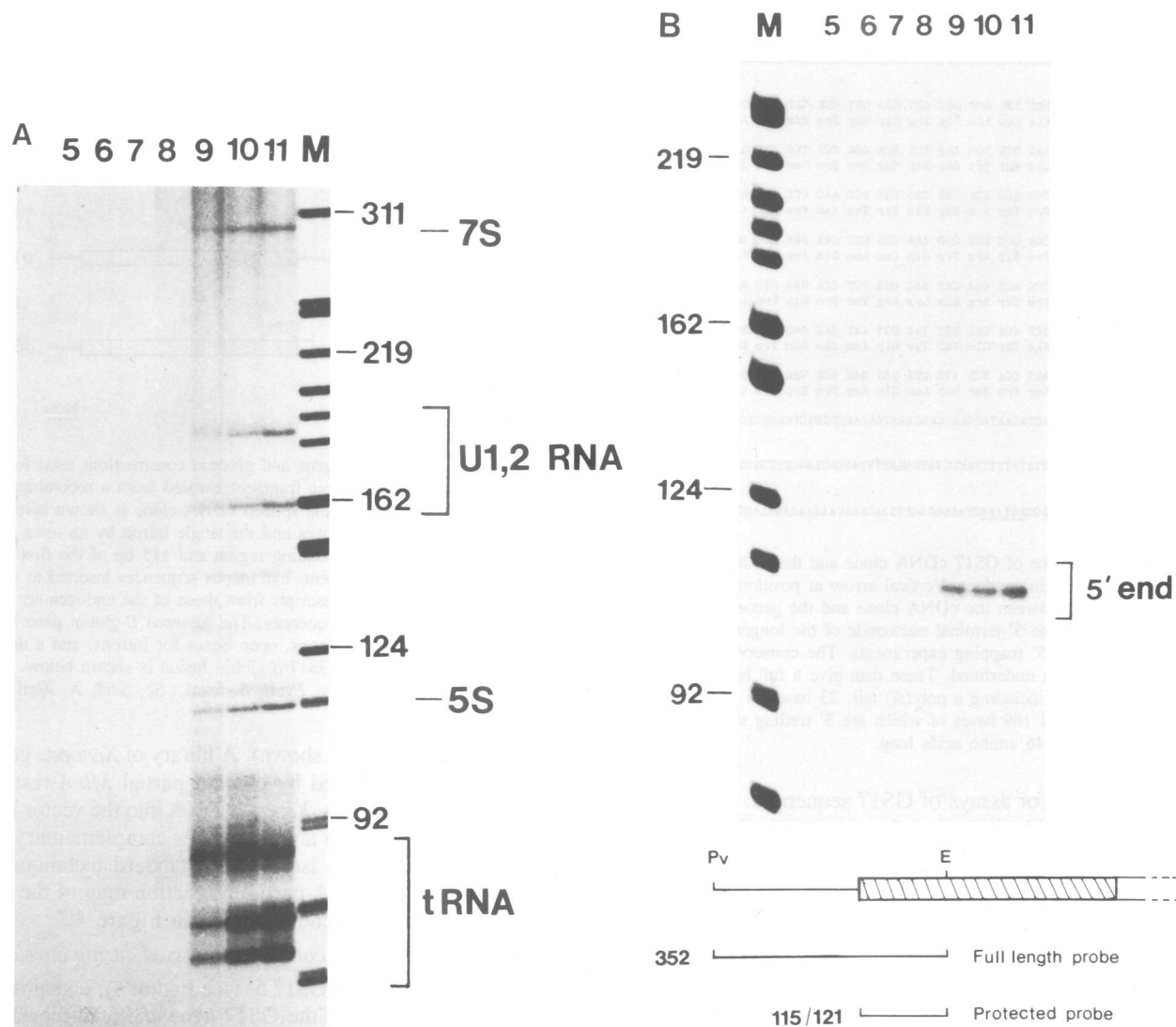


**Fig.1.** Northern blot analysis of GS17 RNA during early frog development. Each track contains the total RNA from five embryos, fractionated on a 1.2% formaldehyde agarose gel. The embryonic stage from which the RNA was isolated is indicated at the top of each track. **A:** shows a 12-h exposure of an autoradiogram of the filter probed with <sup>32</sup>P-labelled insert of the GS17 cDNA clone. The arrow indicates the unspliced RNA (see text). This filter was washed to remove bound probe and rehybridized with a <sup>32</sup>P-labelled insert of a *Xenopus* fibronectin cDNA clone. An 8-h exposure is shown in **B**. Both cDNA inserts were nick-translated to a specific activity of  $6 \times 10^8$  c.p.m./ $\mu$ g. DNA size markers are indicated in **A**. The size of the fibronectin RNA is  $>6$  kb.

none of which appear to contain GS17 RNA. In addition to the 750-base transcript, a faint band corresponding to a transcript of 1200 bases is evident only at stage 10 (Figure 1). The size of this transcript suggests that it is the unspliced pre-mRNA (see below). These results show that GS17 is only transiently expressed during embryogenesis, its transcripts found primarily during gastrulation.

Evidently two factors contribute to the transient appearance of GS17 mRNA. First, as judged by the accumulation of stable RNA, the GS17 gene is transcriptionally active only from mid-blastula to mid-gastrula, a period of  $\sim 5$  h. The fact that the GS17 pre-mRNA is found only at stage 10 supports this argument. Secondly, GS17 mRNA rapidly disappears suggesting that, unlike most mRNAs present in the developing embryo, GS17 mRNA is quite unstable. We estimate that it is degraded with a half-life of 3–4 h.

In addition to the expression pattern for GS17, Figure 1 shows the developmental expression of a typical maternal mRNA, in this case cellular fibronectin. The comparison of the expression profiles of GS17 and fibronectin mRNAs shows that the dramatic changes in GS17 mRNA levels are not due to variable recovery



**Fig. 2.** Onset of GS17 transcription. **(A)** *In vivo* labelling of newly transcribed RNAs. Eggs were injected with  $1 \mu\text{Ci}$  of  $[^{32}\text{P}]\text{UTP}$  at first cleavage and total RNA was extracted at various times and fractionated in a 6% acrylamide-8.3 M urea gel. Each track contains the RNA from three embryos. The time (h) after fertilization at which the RNA was isolated is indicated at the top of each track. The positions of some abundant labelled RNA species are indicated. At this temperature ( $18^\circ\text{C}$ ) new transcripts are first detected at 9 h post-fertilization. **(B)** Detection of GS17 transcripts by RNase mapping. Portions of the same RNA samples analyzed in A above were assayed for the presence of GS17 RNA using a single-stranded SP6 RNA probe. The RNA from one embryo equivalent was tested for each time point and the protected RNA fragments were fractionated on a 6% acrylamide-8.3 M urea gel. The continuously labelled RNA probe and the predicted RNase protection fragments are shown in the diagram below the autoradiogram. M: size markers, labelled *Hpa*II fragments of pBR322 DNA.

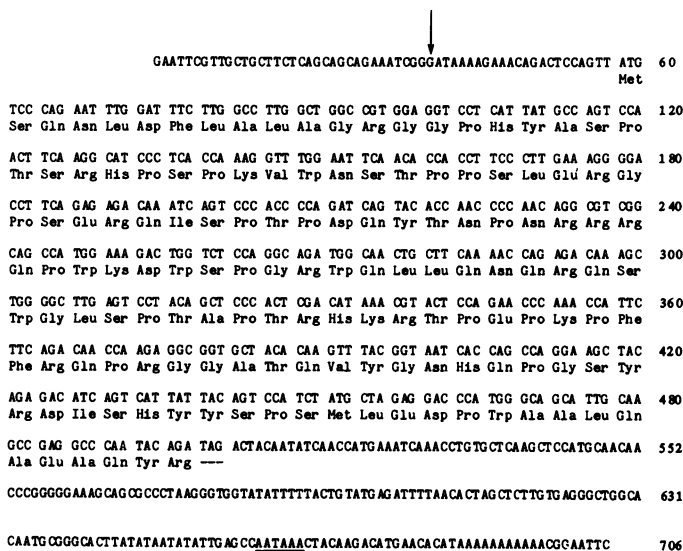
of RNA from the different embryonic stages since the fibronectin RNA remains relatively constant at the early stages when GS17 mRNA first appears. Secondly, while a major increase in the amount of GS17 mRNA occurs by stage 10, no increase is seen in the amount of fibronectin mRNA until stage 12, at least 3 h later. Therefore, although the two genes both become transcriptionally active early in development, they do not appear to be activated simultaneously. This observation implies that GS17 and fibronectin are subject to different mechanisms of control and are not merely activated by some general derepression mechanism that may occur at the MBT (Newport and Kirschner, 1982a, 1982b).

#### *The onset of GS17 transcription*

Genes that are the first to be transcriptionally activated, precisely at the MBT, are more likely to be controlled by maternal fac-

tors than those genes that are activated later in development. While it is clear that tRNAs and snRNAs are some of the first sequences to be transcribed and that new transcription of mRNA has occurred by mid-gastrula (Sargent and Dawid, 1983), it is not known if any mRNAs are transcribed at the MBT. A longer exposure of the blot shown in Figure 1 suggested that GS17 transcripts appear before stage 10, at stage 8. We therefore decided to determine exactly when GS17 transcripts appear relative to the MBT and the onset of general transcription.

To determine the time at which the MBT occurs, all new transcripts in the developing embryo were labelled by microinjection of  $[^{32}\text{P}]\text{UTP}$  directly into the fertilized egg (Newport and Kirschner, 1982a). Starting at the late cleavage state, before the MBT, groups of embryos were collected at 1-h intervals and total RNA was extracted. A sample of the labelled RNA was fractionated on a denaturing acrylamide gel for direct visualization of newly labelled RNA species and another portion of the RNA



**Fig. 3.** Nucleotide sequence of GS17 cDNA clone and the deduced amino acid sequence of the protein product. Vertical arrow at position 35 indicates the point of divergence between the cDNA clone and the genomic clone. This site also represents the 5'-terminal nucleotide of the longer species of GS17 mRNA detected in 5' mapping experiments. The conserved poly(A) addition site, AATAAA is underlined. These data give a full length for the mRNA of 652 bases, not including a poly(A) tail, 23 bases of which are 5' untranslated sequences and 189 bases of which are 3' trailing sequences. The predicted protein is 146 amino acids long.

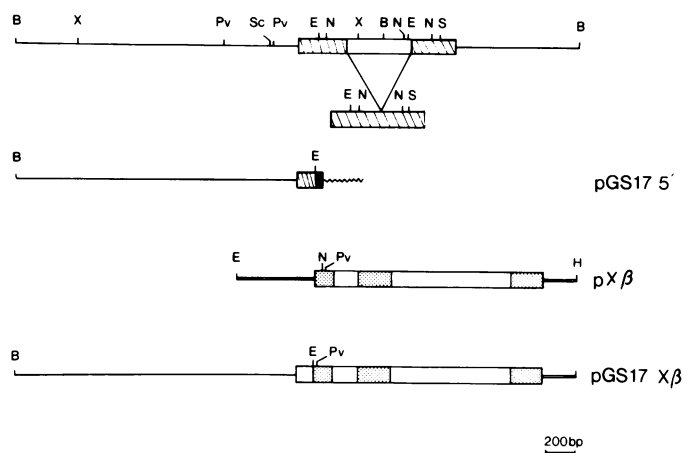
sample was retained for assays of GS17 sequences. The results presented in Figure 2A show that the new transcripts become clearly detectable at 9 h after fertilization, a blastula stage in this experiment. This result is in complete agreement with previous studies which showed that the general onset of transcription occurs at this stage and that the majority of the identifiable sequences transcribed are tRNAs and snRNAs (Newport and Kirschner, 1982a; Busby and Reeder, 1983). GS17 mRNA was detected by 5' end mapping with a highly labelled RNA probe. The results (Figure 2B) show that GS17 sequences are first detected in total RNA extracted from 9-h embryos, precisely the same sample in which newly synthesized RNA was first observed. We conclude, therefore, that the GS17 gene is transcribed at the MBT and is among the very first mRNAs to be transcribed in early embryonic development.

*Nucleotide sequence of a full-length GS17 cDNA clone*

The GS17 cDNA clone was sequenced to characterize further the gene and with the hope that the deduced amino acid sequence might provide some insight into the properties or function of the protein. RNA blotting experiments (Figure 1) show that GS17 was ~750 bases long and the total length of GS17 sequences excised from the  $\lambda$ gt10 recombinant was ~700 bases, nearly a full-length copy of the mRNA. The nucleotide sequence and deduced amino acid sequence of the *Xenopus* GS17 cDNA clone are presented in Figure 3. The protein consists of 146 amino acids and has a predicted mass of 17 000 daltons, hence the name GS17. A computer search of the NIH protein sequence database failed to identify any significant regions of homology between GS17 and other known protein sequences. Two notable features of the protein sequence are a large net basic charge (+ 14) and an unusually high proline content (21 of 146 amino acids).

*Isolation of a genomic clone for GS17*

Southern blots performed under high stringency conditions indicate that a single copy of the GS17 gene is present in the



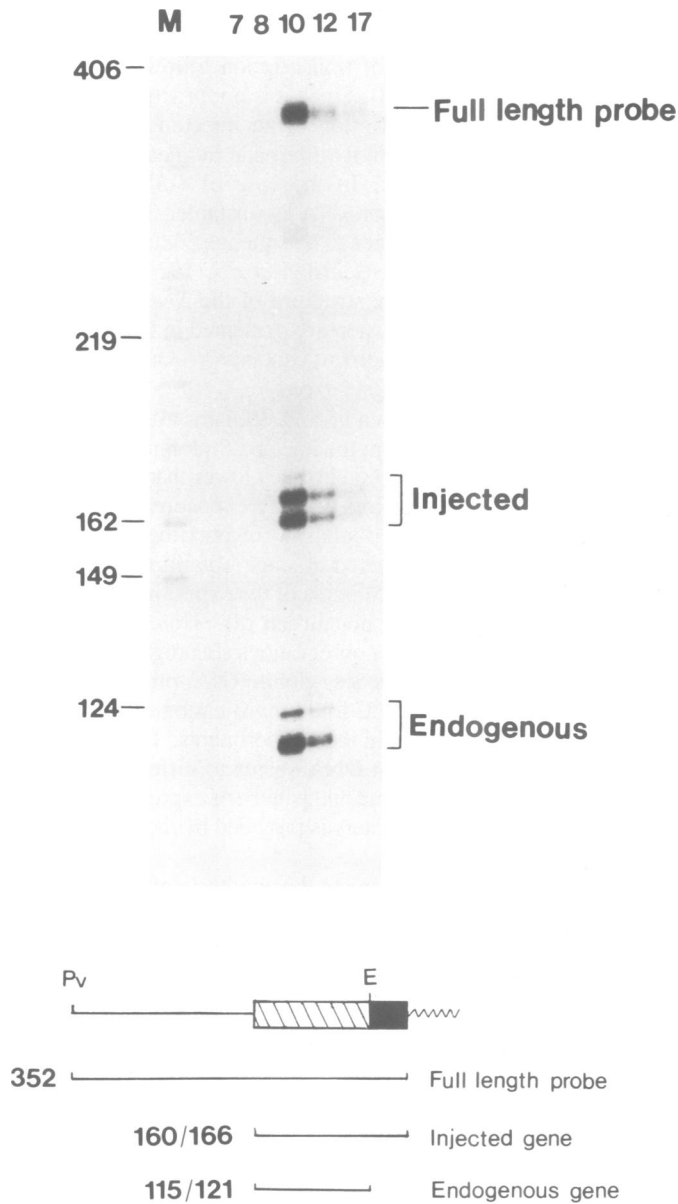
**Fig. 4.** Structure of the GS17 gene and plasmid constructions used for injection studies. A 4-kb genomic fragment isolated from a recombinant  $\lambda$  phage is shown at the top and the spliced cDNA clone is shown below. Exons are shown by hatched boxes and the single intron by an open box. pGS17 5' contains 2 kb of 5'-flanking region and 115 bp of the first exon of GS17. The black box represents extraneous sequences inserted to mark the gene and distinguish its transcripts from those of the endogenous gene.  $\Lambda$  represents pSP64 plasmid sequences. The *Xenopus*  $\beta$ -globin gene is shown with stippled boxes as exons, open boxes for introns, and a double line for flanking regions. The GS17/ $\beta$ -globin fusion is shown below. B, *Bam*HI, E, *Eco*RI, N, *Nco*I; Pv, *Pvu*II; S, *Sma*I.; Sc, *Sac*I; X, *Xba*I.

*Xenopus* genome (data not shown). A library of *Xenopus* genomic sequences was constructed by cloning partial *Mbo*I restriction fragments of high mol. wt. *Xenopus* DNA into the vector EMBL 4 (Frischauf *et al.*, 1983) and sequences complementary to the GS17 cDNA clone were isolated by standard techniques (see Materials and methods). A partial restriction map of the region surrounding the GS17 gene is shown in Figure 4.

*The injected GS17 gene is correctly expressed during development*

A plasmid construction, pGS17 5' (see Figure 4), containing 115 bases of the first exon of the GS17 gene and a further 2 kb of 5'-flanking sequences was injected into developing embryos at the time of the first cell division. The injected gene was marked with a 45-base fragment of foreign DNA so that transcripts could be distinguished from those of the endogenous (chromosomal) GS17 genes. After injection the embryos were allowed to develop and groups of embryos were sampled at various stages of development in order to follow the appearance of transcripts from the GS17 promoter. Total RNA was extracted from embryos at five developmental stages ranging from early blastula through to tail-bud tadpole and transcripts initiated at the correct 5' ends of both the marked injected gene and the endogenous GS17 gene were detected by RNase mapping (Figure 5). Initiation from the endogenous gene protects two RNA fragments, 115 and 121 bases long. This doublet probably corresponds to two transcription initiation sites, though we have not ruled out processing of the RNA transcripts. Transcripts from the marked gene also protect two RNAs, 160 and 166 bases long.

The results presented in Figure 5 show that prior to mid-blastula no transcription is observed from either the endogenous GS17 promoter or from the injected GS17 5' DNA. At subsequent stages of development the expression of the marked gene parallels the expression of the endogenous GS17 gene. In both cases, transcripts begin to accumulate at mid-blastula and reach peak levels at the mid-gastrula stage. Though these assays do not directly test transcription rates, the results indicate that both the injected and the endogenous GS17 promoters are only transcriptionally



**Fig. 5.** Expression of GS17 DNA in injected embryos. 50 pg of supercoiled GS17 5' DNA was injected into fertilized eggs as the first cleavage began. Injected eggs were allowed to develop to the stage indicated at the top of each track and total RNA was extracted. Transcripts from the endogenous and injected GS17 sequences were detected by RNase mapping with the same single-stranded SP6 RNA probe. This continuously  $^{32}\text{P}$ -labelled RNA probe and the RNase-protected fragments that map the 5' ends of GS17 RNAs are shown below the autoradiogram. Total RNA from one embryo equivalent was assayed and the protected fragments were fractionated in a 6% acrylamide-8.3 M urea gel. The structure of the pGS17 5' gene is shown more completely in Figure 4. The size markers, M, are *Hpa*II fragments of pBR322.

active from the mid-blastula transition until about midway through gastrulation. Thereafter, transcripts from both the injected and the chromosomal GS17 genes disappear and are absent after neurulation.

This and other experiments have shown that while transcripts from the injected GS17 DNA disappear during neurulation, they do so more slowly than their natural counterpart. As discussed earlier, the disappearance of natural GS17 mRNA from embryos suggests that the mRNA is specifically degraded during neurula-

tion. It is not obvious why transcripts from the marked GS17 genes should also disappear. However, since pGS17 5' does not contain the complete genomic sequence of GS17, the transcripts will terminate at sequences in the plasmid vector and therefore lack a proper 3' end with a poly(A) tail. These transcripts, containing heterogeneous 3' portions of bacterial sequences, are apparently unstable in the developing frog embryo. If a stable transcript had been produced from the marked gene we would have expected to observe an increase in the amount of transcript while the GS17 promoter was active and then a constant high level of RNA at subsequent developmental stages. An example of this type is presented in the GS17/ $\beta$ -globin fusion experiments described below.

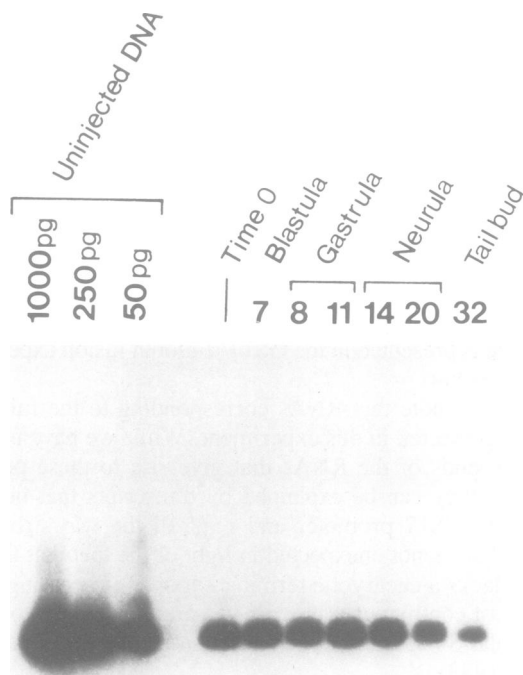
Finally, we note that RNAs corresponding to the full-length probe are protected in this experiment. While we have not mapped the 5' ends of the RNAs that give rise to these protected fragments, they can be explained by transcripts that initiate at the injected GS17 promoter and read all the way around the plasmid. This is not unexpected in light of the fact that this construction lacks a eucaryotic termination or 3' processing signal. This type of continuous transcript has previously been observed with circular SV40 transcripts injected into oocytes (Wickens and Gurdon, 1983).

#### Fate of injected GS17 DNA

Previous experiments in which plasmid DNA was injected into developing embryos showed varying degrees of amplification of the injected DNA. This was first studied by Harland and Laskey (1980) who showed that a small portion of DNA injected into fertilized eggs is correctly replicated through the first few cell divisions. Several groups have investigated the persistence of the injected DNA through later developmental stages. For example, *Xenopus* and rabbit globin genes and sea urchin histone genes were amplified 50- to 200-fold following injection (Bendig and Williams, 1983; Rusconi and Schaffner, 1981; Etkin *et al.*, 1984). In these cases, the injected plasmid DNA reached peak levels at late blastula to mid-gastrula and then most (up to 95%) of the DNA disappeared by the tadpole stage. In contrast, injected yeast tRNA genes (Newport and Kirschner, 1982b) and *Xenopus* rRNA genes (Busby and Reeder, 1983) were not amplified at all, and injected SV40-CAT genes were only amplified ~5-fold (Etkin *et al.*, 1984).

The degree to which the injected GS17 DNA is amplified is significant for two reasons. First, a large amplification factor greatly increases the amount of transcription template present in the embryo. Secondly, dramatic changes in the amount of DNA template would affect the interpretation of the expression pattern of the injected genes. In principle, a rapid increase in template DNA up to mid-gastrula followed by a rapid decrease in template DNA levels as development progressed could account for the expression profile shown in Figure 5. To test this possibility, we determined the extent to which injected GS17 DNA is amplified. Total nucleic acid from the equivalent of one injected embryo was digested with *Eco*RI and used for a Southern blot. The results (Figure 6) show that the amount of injected DNA remains constant from the fertilized egg through neurula, the period during which GS17 is expressed. After neurulation, there is a decrease in the amount of injected DNA, but this is long after GS17 expression has stopped. Therefore, changes in the amount of GS17 template cannot account for the expression profile observed in Figure 5.

In a parallel series of experiment we have assayed the state of the injected DNA using Southern blots. We find that the in-



**Fig. 6.** Persistence of injected DNA during early development. Fertilized eggs were injected with 50 pg of supercoiled pGS17 5' DNA (Figure 4) and allowed to develop to the stage indicated at the top of each track. Total nucleic acid was extracted and the DNA from one embryo equivalent was digested with *EcoRI*, fractionated in a 0.8% agarose gel and blotted to nitrocellulose. The filter was hybridized with  $^{32}\text{P}$ -labelled vector DNA. The time 0 track contains DNA extracted from embryos immediately following injection. *EcoRI* digested, uninjected DNA standards of pGS17 5' DNA were included in the gel to estimate the degree to which the injected DNAs were amplified.

jected DNA remains extrachromosomal in a supercoiled form, i.e., there is no indication that the GS17 DNA integrates into the host chromosome up until the tailbud stage (data not shown). We cannot rule out a very low level of integration, but our results show that >99% of the injected template is extrachromosomal.

We conclude that the amount of injected GS17 template remains constant throughout early development and that the transient expression of the injected DNA (Figure 5) reflects changes in the activity of the GS17 promoter. Moreover, we find that most of the injected DNA is not integrated into the host genome and persists as an extrachromosomal element.

#### *Transcription of injected Xenopus globin and GS17/globin fusion genes in embryos*

As noted in the Introduction, several experiments have examined the expression of genes injected into *Xenopus* embryos and in some cases a low level of transcription has been shown to start soon after the MBT (Bendig, 1981; Rusconi and Shaffner, 1981; Bending and Williams, 1983, 1984; Etkin *et al.*, 1984; Etkin and Balcells, 1985). In the light of these results, how can we be sure that transcription of injected GS17 genes is not a similar non-specific event? To address this issue we have compared the

expression of the injected GS17 promoter with the expression of another promoter, the *Xenopus* adult  $\beta$ -globin gene, previously shown to give a low level of transcription following the MBT (Bendig and Williams, 1983). Since it is not practical to directly measure the rate of transcription of an injected promoter, inferences about promoter activity are made by measuring the accumulation of stable RNA. In the case of GS17 this is not straightforward because the mRNA is unstable. Therefore, we have fused the GS17 promoter to a sequence coding for a stable transcript, a globin mRNA (Gurdon *et al.*, 1974; Bendig and Williams, 1983, 1984). The structure of the *Xenopus*  $\beta$ -globin gene and the GS17/globin fusion are presented in Figure 4. Note that the GS17 'promoter' fragment contains 2.0 kb of 5'-flanking region and 115 bp of the first exon.

As previously shown by Bendig and Williams (1983), we find a low level of transcription from the injected  $\beta$ -globin gene after the MBT. This level of transcription is much lower than that observed for injected GS17 and requires a 7-day exposure of the gel for detection (Figure 7A). Faint bands corresponding to the 5' end of globin mRNA are first detected at stage 10 and these continue to accumulate throughout the course of the experiment, indicating that the globin promoter is not turned off. This pattern of expression agrees with previous observations (Bendig and Williams, 1983), except that we observe less globin RNA, probably because we do not find the 50- to 100-fold amplification of the injected template that they observed in their experiments. Thus, the transcription of injected  $\beta$ -globin DNA seems to differ from that of GS17 DNA in both the amount and pattern of expression. A direct comparison of the two promoters is provided by the results shown in Figure 7B.

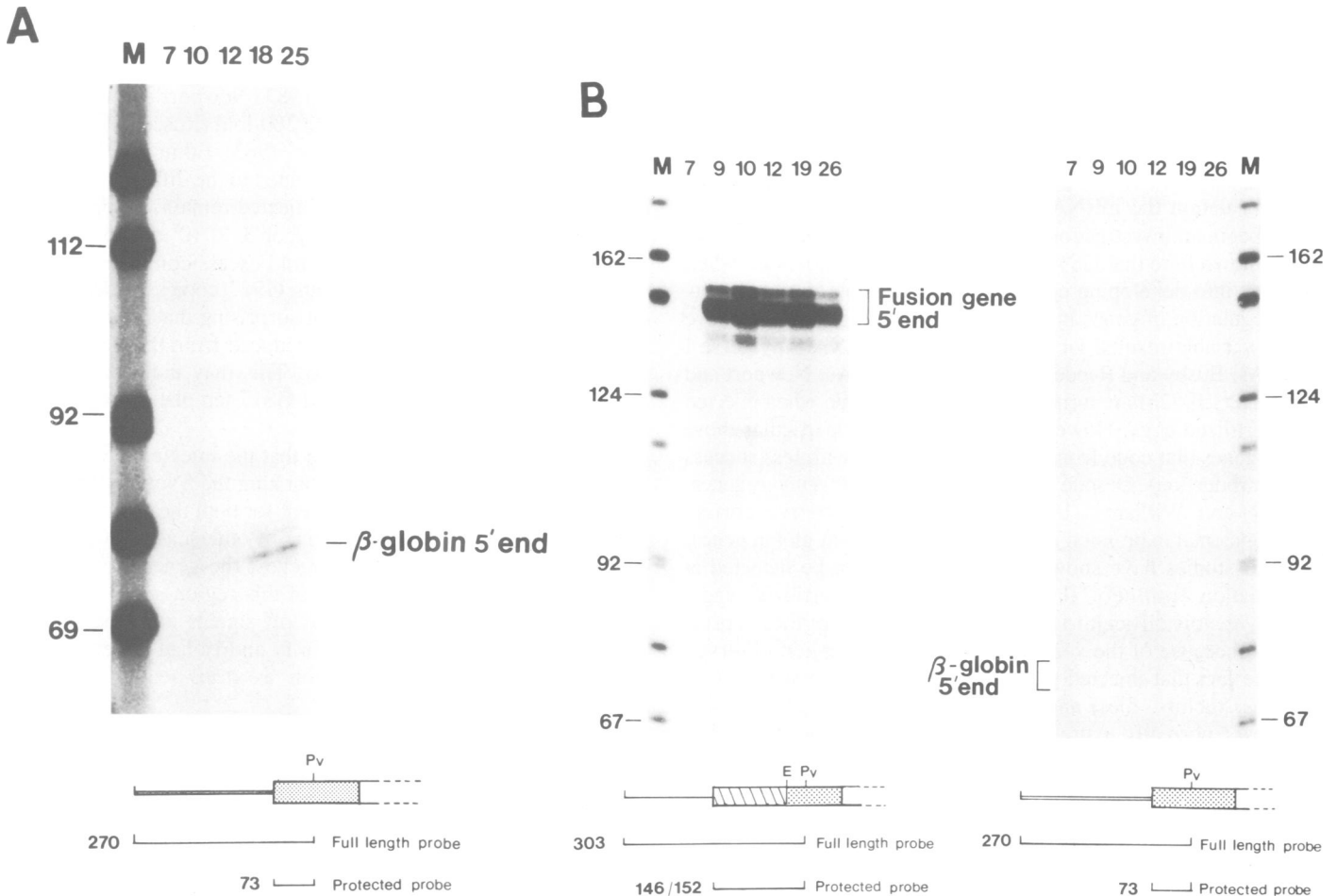
The GS17/globin DNA directs the synthesis of high levels of transcripts in injected embryos (Figure 7B). A 6-h exposure shows a conspicuous band at a size predicted for the 5' end of the fusion transcript. Transcripts are not detected prior to the mid-blastula transition, but by stage 10 they are abundant and by stage 12, mid-gastrula, the transcripts have accumulated to peak levels. After mid-gastrula the fusion transcript persists at an approximately constant level throughout the developmental stages tested. In marked contrast, transcripts from the  $\beta$ -globin gene are below the level of detection in a 6-h exposure (Figure 7B).

The expression pattern observed from the fusion gene is expected for a GS17 promoter that is transiently active from the mid-blastula to mid-gastrula stage and that directs the synthesis of a stable transcript. Unlike the natural case where GS17 mRNA rapidly disappears, the fusion transcript persists through later development, probably because it contains *Xenopus*  $\beta$ -globin mRNA sequences that are quite stable in injected embryos. If the injected GS17 promoter remained transcriptionally active during late stages, one would have expected to see a continuous increase in the amount of the fusion transcript. Since that is not observed, this fusion experiment supports the conclusion drawn from injecting GS17 5' DNA (Figure 5), namely that the injected GS17 promoter is turned on and off during early development.

As a control, the amount of template was assayed as before (Figure 6) and the results (not shown) reveal that equal amounts of both templates were present at each developmental stage. In this case, both templates were amplified 3- to 5-fold, though in other experiments using the same templates amplification was not observed. The important point is that the injected globin and fusion templates are present in equal amounts so that the relative activities of the two promoters is directly comparable.

Quantitation of the results presented in Figure 7B shows that





**Fig. 7.** Expression of  $\beta$ -globin and GS17/ $\beta$ -globin fusion genes in injected embryos. 50 pg of supercoiled DNA was injected into fertilized eggs at the first cleavage. The structures of the injected DNAs are shown more completely in Figure 4. Injected eggs were allowed to develop to the stage indicated at the top of each track and total RNA was extracted. RNase mapping of the 5' ends of the transcripts was performed with RNA from one embryo equivalent using the single-stranded SP6 RNA probes shown below each autoradiogram. The protected fragments were fractionated in a 6% acrylamide-8.3 M urea gel. **(A)** 7-day autoradiographic exposure of probe protected by RNA extracted from embryos injected with *Xenopus*  $\beta$ -globin gene. **(B)** 6 h autoradiographic exposure of probe protected by RNA extracted from embryos injected with either GS17/ $\beta$ -globin fusion gene or the  $\beta$ -globin gene alone.

the injected GS17 promoter is ~200- to 500-fold more active than the injected globin gene promoter in injected embryos. We conclude that transcription of the injected GS17 gene is correctly regulated and is not the result of a non-specific transcription of injected DNA. Moreover, in conjunction with the results shown in Figure 5, these experiments suggest that the *cis*-acting signals that direct RNA synthesis of GS17 during gastrulation are contained within the 5' region of the GS17 gene.

### Discussion

We show here that GS17 is among the very first genes to be expressed by the embryonic genome, exactly at the MBT. GS17 transcripts appear at the same time as new tRNA and snRNA transcripts and considerably earlier than rRNAs (Shiokawa *et al.*, 1981; Busby and Reeder, 1983) or mRNAs such as muscle actin (Mohun *et al.*, 1984) or fibronectin (this paper). It is not understood how selective gene transcription at the mid-blastula transition is accomplished, but the very first genes to be transcribed are probably controlled by maternal factors. A dependence on controlling factors expressed by the maternal genome during oogenesis is inferred from the fact that the embryonic genome is transcriptionally inactive prior to the MBT. Thus, if proteins

are involved in the transcriptional activation of genes at the MBT, those proteins must be provided directly from a maternal pool or indirectly by translation of maternal mRNAs. In line with this reasoning, we assume that all the factors involved in regulating the transcription of a gene like GS17 must be present, in an active or inactive form, in the egg. It is perhaps worth noting that this may be quite different from the situation in terminally differentiated cells. In this latter case the activation of a gene may be the end result of a long series of regulatory events with many of the key factors present only in precursor cells that are difficult or impossible to obtain.

An intriguing property of GS17 expression is the fact that its transcripts are found only during a brief period of development, principally through gastrulation. Sargent and Dawid (1983) have also identified genes that are expressed early, at gastrulation, and some of these are not expressed at later embryonic stages. In the case of GS17, our studies indicate that the transient expression of the gene is the result of a limited period of transcription coupled with the instability of the transcript. The level of GS17 mRNA reaches its maximum only 2 h after transcription is initiated and the unprocessed pre-mRNA is observed only at stage 10 gastrula. These results suggest that transcription from the GS17 promoter stops ~4–5 h after it is first initiated. Unlike most

mRNAs in the developing embryo which are extremely stable (Colot and Rosbash, 1982), the GS17 mRNA disappears with a half-life of  $\sim 3-4$  h. It is formally possible that the disappearance of GS17 mRNA after gastrulation is the result of continuous transcription combined with an increased rate of mRNA degradation. However, we believe the simplest interpretation of our results is the GS17 promoter is active transiently and that by neurulation the mRNA is degraded.

To begin an investigation of this transcriptional regulation, we have shown here that GS17 genes are correctly expressed when injected into developing eggs. Previous attempts to obtain correct regulation of genes injected into developing frog eggs have given variable results. Genes transcribed by RNA polymerase I (rRNAs, Busby and Reeder, 1983) and III (tRNAs, Newport and Kirschner, 1982b) are turned on at the proper time when injected into fertilized eggs. However, other studies, in particular those using genes that code from mRNAs, have met with less success (see Introduction). Despite the use of homologous *Xenopus* genes, Bendig and Williams (1983, 1984) did not observe correct developmental expression of either tadpole or adult globin genes. Several studies have shown that transcripts can be detected at gastrulation from Pol II genes injected into fertilized eggs. However, it is difficult to assess the significance of these observations because of the very low level of transcription observed and the fact that injected genes were from foreign species (sea urchins, rabbits, flies) and/or viruses (adenovirus, SV40) that are never normally expressed during frog embryogenesis. The one conclusion that can be drawn from these studies is that there can be a low level of transcription after the MBT and that this low level may continue throughout embryogenesis as if there were some 'leakage' due to the vast amounts of DNA injected.

In the light of these previous results, it is important to show that the expression pattern of the injected gene reflects normal developmental regulation. With respect to the transcriptional activation of the injected DNA, several results suggest that this step is correctly regulated. First, transcripts from the injected GS17 DNA begin to accumulate at the same time as those from the endogenous gene, shortly after the MBT. Secondly, we note that the amount of transcript synthesized from the injected DNA is about the same as the endogenous GS17 mRNA. Thirdly, we have compared the activity of the GS17 promoter with that of the *Xenopus*  $\beta$ -globin gene and find that the GS17 promoter is 200–500 times more active than the globin promoter at the MBT. Thus, the transcriptional activation of injected GS17 DNA is not merely the result of some general low level of transcription that occurs on all injected DNAs.

Our results also suggest that the GS17 promoter is turned off during gastrulation. Transcripts from the injected GS17 DNA disappear during neurulation like those of the endogenous gene. If the injected DNA remained transcriptionally active after gastrulation we should have observed a continual increase in the amount of GS17 transcripts. Moreover, when the GS17 promoter elements are fused to the body of a  $\beta$ -globin gene in order to direct the synthesis of a stable RNA, this fusion transcript does not continually increase in amount during later developmental stages. In contrast, some injected DNAs (e.g.,  $\beta$ -globin, Figure 7a) are turned on and remain on so that there is a steady increase in the amount of transcript accumulated. Together, these results and those obtained with Pol I and III genes (Busby and Reeder, 1983; Newport and Kirschner, 1982b) suggest that genes that are normally expressed early in embryonic development may be correctly regulated when injected into fertilized eggs.

In most of our experiments the injected DNA is not detectably

amplified, though in some cases we do observe a low level ( $< 5$ -fold) of amplification. Previous reports reveal a wide range in the degree to which injected DNAs are amplified, from practically none (Busby and Reeder, 1983; Newport and Kirschner, 1982b; Etkin *et al.*, 1984) to 50- to 200-fold (Rusconi and Schaffner, 1981; Bendig and Williams, 1983; Etkin *et al.*, 1984). Whether this variation can be attributed to the different templates used or to the amounts of DNA injected remains an open question. In our experiments,  $\sim 50$  pg or  $3 \times 10^7$  copies of GS17 were injected, which is an  $\sim 10^3$ -fold excess compared with the calculated number of endogenous GS17 copies present at the MBT. In this light it is somewhat surprising that approximately equal amounts of transcript are produced from the injected and endogenous templates (Figure 5). This may indicate that only a small proportion of the injected GS17 templates are active in transcription.

Finally, it is interesting to note that the injected GS17 DNA, which contains  $\sim 2$  kb of DNA flanking the 5' end of the gene, has the *cis*-acting elements sufficient for both the correct switch-on and switch-off of transcription. In other cases, such as the ribosomal genes and the tRNA genes, the genes are activated permanently. Further dissection of this region should allow us to determine whether the on and off signals are separable. In addition, we may be able to identify and isolate maternal components from eggs that function as *trans*-acting factors in regulating GS17 expression.

## Materials and methods

### Materials

*X. laevis* were purchased from Xenopus I, Ann Arbor, MI and fed chopped liver twice per week. Radionucleotides were purchased from Amersham, AMV reverse transcriptase from Life Sciences, Inc., SP6 RNA polymerase from Promega Biotec and all restriction enzymes from Promega Biotec and New England Biolabs.

### Injection of fertilized eggs

Eggs were obtained by injecting *Xenopus* females with mare serum gonadotropin (500 units) followed 2–5 days later by the injection of human chorionic gonadotropin (1000 units). Eggs were laid in HSB (120 mM NaCl, 20 mM Tris-HCl, pH 7.6, 1 mM KCl, 1 mM MgSO<sub>4</sub>, 2 mM NaHCO<sub>3</sub>) and fertilized by the addition of sperm from macerated testis in tap water. About 20 min after fertilization eggs were dejellied in 3% cysteine (pH 7.5) and transferred to 0.1 X MSBH (Gurdon, 1976), 5% Ficoll (recommended by J. Gerhardt). Injections of 20 nl into the vegetal hemisphere were performed as the first cleavage furrow appeared at the animal pole. Eggs were incubated in 0.1 X MBSH, 5% Ficoll either at room temperature or at 18°C. All embryonic stages given are according to Nieuwkoop and Faber (1967).

### Construction and screening of a *Xenopus* gastrula cDNA library

A cDNA library was constructed in  $\lambda$ gt10 (Huyhn *et al.*, 1985), using poly(A)<sup>+</sup> RNA isolated from stage 11 gastrula embryos. The details of cDNA synthesis and insertion into the vector are described elsewhere (Rebagliati *et al.*, 1985). The library contains  $\sim 10^6$  independent recombinants with an average insert size of 1.1 kb.

The gastrula cDNA library was plated on *Escherichia coli* C600/hfl at a density of 2000–3000 plaques per 15 cm plate. Recombinants were transferred to nitrocellulose filters in duplicate (Benton and Davis, 1977), denatured *in situ* by autoclaving (G. Struhl, personal communication; Weeks *et al.*, 1985), baked for 1 h *in vacuo* at 80°C and pre-hybridized in 50% formamide at 42°C as described (Maniatis *et al.*, 1982). Hybridizations were carried out in the same buffer with  $9 \times 10^6$  c.p.m. per filter using <sup>32</sup>P-labelled cDNA prepared from either defolliculated oocyte or stage 11 poly(A)<sup>+</sup> RNA. The filters were washed in 2 X SSC, 0.1% SDS at 65°C and exposed with intensifying screens at  $-70^\circ\text{C}$  for 3 days.

Clones yielding a differential response were purified and subjected to two further tests. First, a large number of the differential clones were found to contain mitochondrial sequences as judged by hybridization to a *Xenopus* mitochondrial DNA clone (Rastl and Dawid, 1979). *Bona fide* differential clones were selected from the remaining candidates by Northern blot tests. 1  $\mu\text{g}$  of poly(A)<sup>+</sup> RNA from defolliculated oocytes and stage 11 embryos was fractionated in adjacent tracks on a denaturing agarose gel, transferred to nitrocellulose, and hybridized with <sup>32</sup>P-labelled insert isolated from putative differential clones. These Northern



blots provided a sensitive test for sequences entirely absent from oocyte RNA, but present in gastrula RNA.

#### Isolation of *Xenopus fibronectin* cDNA clones

$8 \times 10^5$   $\lambda$ gt10 recombinants from an oocyte cDNA library (Rebagliati *et al.*, 1985) were screened under low stringency hybridization conditions (McGinnis *et al.*, 1984) for sequences homologous to rat fibronectin using  $^{32}$ P-labelled rfl1 (Schwarzbauer *et al.*, 1983). The 48 positive clones were assayed for insert size. The longest cDNA insert contains two *Eco*RI fragments, 2.6 and 0.4 kb in length, and these were subcloned into pSP65 to generate pXFA and pXFB, respectively. DNA sequencing of a 450-bp region near the 5' end of pXFA predicts an amino acid sequence that is 65% homologous to the rat plasma fibronectin sequence, between nucleotides 316 and 798 as reported by Schwarzbauer *et al.* (1983). Sequence comparisons suggest that this *Xenopus* clone encodes a cellular fibronectin (R. Hynes, personal communication). pXFB contains sequences extending further towards the 5' end of the mRNA.

#### Construction and screening of a *Xenopus* genomic library

DNA was isolated from st 33 embryos obtained from two separate matings. The DNA fragments resulting from *Mbol* partial digestion were fractionated by sucrose density gradients and purified before insertion into the *Bam*HI site of  $\lambda$ EMBL4 (Frischauf *et al.*, 1983). After *in vitro* packaging,  $10^6$  recombinants were plated on *E. coli* K802 and amplified. Digestion of 18 randomly selected clones with *Eco*RI shows that the average insert size is ~ 16 kb.

To isolate a genomic copy of GS17 DNA,  $8 \times 10^5$  recombinants from this library were screened with  $^{32}$ P-labelled insert from the GS17 cDNA. 24 homologous clones were identified and the partial restriction map of one clone,  $\lambda$ X GS17.1 is shown in Figure 4. Restriction fragments of the genomic clone were subcloned into pSP64 (Melton *et al.*, 1984) for further analysis.

#### DNA sequencing

The DNA sequence of the GS17 cDNA clone and portions of the fibronectin cDNA clone were obtained using the chain termination method of Sanger *et al.* (1977). Restriction fragments were cloned into M13 vectors (Messing and Vieira, 1982) and in all cases the sequence of both strands was determined.

#### Plasmid constructions

For injection studies, it is necessary to distinguish transcripts of the injected GS17 DNA from those of the endogenous gene. A marked GS17 gene that contains most of the 5'-flanking region was constructed as follows. The 2.6-kb *Bam*HI fragment of the genomic GS17 DNA was subcloned into *Bam*HI-cut pSP64. Digestion of this DNA with *Eco*RI yields a 2.1-kb fragment which was cloned into *Eco*RI-cut pSP64 to give pGS17 5' (see Figure 4). The polylinker sequences downstream from the *Eco*RI site in the first exon of GS17 act as the marker sequence, shown as a black box in Figure 4.

The *Xenopus*  $\beta$ -globin gene, pX $\beta$ , consists of the 2.4-kb *Eco*RI to *Hind*III fragment of the major adult  $\beta$ -globin gene (Patient *et al.*, 1980), inserted into pSP64.

The fusion gene, pGS17/X $\beta$  was constructed as follows. pX $\beta$  DNA was cleaved at the unique *Nco*I site located at the initiation codon in the first exon of the globin gene. The overhanging ends were filled in with reverse transcriptase and an *Eco*RI 12-mer linker was ligated to the blunt end. Following digestion with *Eco*RI, the 2.1-kb GS17 promoter fragment of pGS17 5' was inserted into the plasmid in place of the  $\beta$ -globin promoter. The 12-mer *Eco*RI linker was used to create the fusion gene so that an open reading frame would be retained in the RNA transcripts.

#### Analysis of transcripts

RNAs were isolated from eggs and embryos by proteinase K digestion, phenol:chloroform extraction, and ethanol precipitation as described previously (Melton and Cortese, 1981). The 5' ends of gene transcripts were mapped with an RNase protection assay using SP6 single-stranded RNA probes as described previously (Zinn *et al.*, 1983; Melton *et al.*, 1984). In each case, appropriate restriction fragments spanning the 5' end of the gene were cloned downstream from an SP6 promoter in pSP64 or pSP65. Uniformly labelled RNA probes ( $6 \times 10^8$  c.p.m./ $\mu$ g) were synthesized from linear templates using SP6 RNA polymerase with [ $\alpha$ - $^{32}$ P]UTP as described previously (Melton *et al.*, 1984). The sizes of full-length probe and the expected RNase A and T1 protected fragments are shown in each figure.

#### Acknowledgements

We thank C. Kintner, R. Myers, T. Maniatis and G. Struhl for comments on the manuscripts and K. Breakey for preparing figures. We thank J. Tankum and R. Hynes for sharing clones and information about rat fibronectin genes. P.K. gratefully acknowledges support from the CSIRO and the Fogarty International Center. This work was supported by grants from the NIH and the Searle Scholars Program.

#### References

- Bendig, M.M. (1981) *Nature*, **292**, 65-67.  
 Bendig, M.M. and Williams, J.G. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6197-6201.  
 Bendig, M.M. and Williams, J.G. (1984) *Mol. Cell. Biol.*, **4**, 567-570.  
 Benton, W.D. and Davis, R.W. (1977) *Science (Wash.)*, **196**, 180-182.  
 Brown, D.D. and Littna, E. (1964) *J. Mol. Biol.*, **8**, 669-687.  
 Busby, S.J. and Reeder, R.H. (1983) *Cell*, **34**, 989-996.  
 Chase, J.W. and Dawid, I. (1972) *Dev. Biol.*, **27**, 504-518.  
 Colot, H.V. and Rosbash, M. (1982) *Dev. Biol.*, **94**, 79-86.  
 Dworkin, M.B. and Dawid, I. (1980) *Dev. Biol.*, **76**, 449-464.  
 Etkin, L.D. and Balcells, S. (1985) *Dev. Biol.*, **108**, 173-178.  
 Etkin, L.D., Pearman, B., Roberts, M. and Bektesh, S.L. (1984) *Differentiation*, **26**, 194-202.  
 Frischauf, A., Lehrach, H., Poustka, A. and Murray, N. (1983) *J. Mol. Biol.*, **170**, 827-842.  
 Gerhart, J.C. (1980) in Goldberger, R. (ed.), *Biological Regulation and Development*, Vol. 2, Plenum Press, NY, pp. 133-316.  
 Goldberg, D.A., Posakony, J. and Maniatis, T. (1983) *Cell*, **34**, 59-73.  
 Gurdon, J.B. (1976) *J. Embryol. Exp. Morphol.*, **36**, 523-540.  
 Gurdon, J.B., Woodland, H.R. and Lingrel, J.B. (1974) *Dev. Biol.*, **39**, 125-133.  
 Harland, R.M. and Laskey, R.A. (1980) *Cell*, **21**, 761-771.  
 Huynh, T., Young, R.A. and Davis, R.W. (1985) in Glover, D.M. (ed.), *DNA Cloning, A Practical Approach*, IRL Press, Oxford and Washington, DC, pp. 49-78.  
 Krumlauf, R., Hammer, R.E., Tilghman, S.M. and Brinster, R.L. (1985) *Mol. Cell. Biol.*, **5**, 1639-1648.  
 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.  
 McGinnis, W., Levine, M., Hafen, E., Kuroiwa, A. and Gehring, W.J. (1984) *Nature*, **308**, 428-433.  
 Messing, J. and Vieira, J. (1982) *Gene*, **19**, 269-272.  
 Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035-7056.  
 Mohun, T.J., Brennan, S., Dathan, N., Fairman, S. and Gurdon, J.B. (1984) *Nature*, **311**, 716-721.  
 Newport, J. and Kirshner, M. (1982a) *Cell*, **30**, 675-686.  
 Newport, J. and Kirshner, M. (1982b) *Cell*, **30**, 687-696.  
 Nieuwkoop, P.D. and Faber, J. (1967) *Normal Table of Xenopus laevis (Daudin)*, published by North Holland, Amsterdam.  
 Palmiter, R.D. and Brinster, R.L. (1985) *Cell*, **41**, 343-345.  
 Patient, R.K., Elkington, J.A., Kay, R.M. and Williams, J.G. (1980) *Cell*, **21**, 565-573.  
 Rastl, E. and Dawid, I.B. (1979) *Cell*, **18**, 501-510.  
 Rebagliati, M.R., Weeks, D.L., Harvey, R.P. and Melton, D.A. (1985) *Cell*, **42**, 769-777.  
 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. natl. Acad. Sci. USA*, **74**, 5436-5467.  
 Sargent, T.D. and Dawid, I.B. (1983) *Science (Wash.)*, **222**, 135-139.  
 Schwarzbauer, J.E., Tamkun, J.W., Lemischka, I. and Hynes, R.O. (1983) *Cell*, **35**, 421-431.  
 Shiokawa, K., Misumi, Y. and Yamana, K. (1981) *Dev. Growth Differ.*, **23**, 579-587.  
 Signoret, J. and Lefresne, J. (1971) *Ann. Embryogen. Morphogen.*, **4**, 113-123.  
 Spradling, A.C. and Rubin, G.M. (1983) *Cell*, **34**, 47-57.  
 Young, P.G. and Zimmerman, A.M. (1973) *Dev. Biol.*, **33**, 196-205.  
 Wickens, M.P. and Gurdon, J.B. (1983) *J. Mol. Biol.*, **163**, 1-26.  
 Zinn, K., DiMaio, D. and Maniatis, T. (1983) *Cell*, **34**, 865-879.

Received on 9 October 1985