Cell, Vol. 39, 479–489, December 1984 (Part 2), Copyright © 1984 by MIT 0092-8674/84/130479-11 $02.00/0

Xenopus 5S Gene Transcription Factor, TFIIIA: Characterization of a cDNA Clone and Measurement of RNA Levels throughout Development

Ann M. Ginsberg, Balas O. King, and Robert G. Roeder
Laboratory of Biochemistry and Molecular Biology
The Rockefeller University
New York, New York 10021

Summary

Initiation of 5S RNA gene transcription in Xenopus oocytes requires a 38,500 dalton polypeptide, TFIIIA. The levels of both 5S RNA and TFIIIA are regulated throughout oogenesis and embryonic development. To delineate the mechanisms by which the corresponding genes are regulated, as well as to determine the primary structure of TFIIIA, we have isolated a cDNA clone that encodes TFIIIA. Using the cDNA clone, we have determined that there is (are) one or a small number of TFIIIA gene(s) per Xenopus haploid genome, and we have estimated the size and levels of TFIIIA RNA throughout Xenopus development. We report sequence homologies between TFIIIA cDNA and regulatory regions of the Xenopus tmet and 5S RNA genes. Implications of these data for developmental regulation of the TFIIIA and 5S RNA genes are discussed.

Introduction

Xenopus laevis contains two types of 5S RNA genes (oocyte-type and somatic-type). These genes are differentially regulated throughout development (reviewed in Korn, 1982). They therefore provide an excellent system for studying developmental regulation of gene expression as well as basic aspects of transcription mechanisms. The transcription of 5S RNA genes requires at least three factors in addition to RNA polymerase III (Segall et al., 1980; Shastry et al., 1982). One of these factors, designated TFIIIA, is specifically required for initiation of transcription of the 5S RNA genes, but is not required for the synthesis of any other polymerase III transcript (e.g., tRNAs, adenovirus VA RNAs) (Engelke et al., 1980; Segall et al., 1980; Shastry et al., 1982). TFIIIA has been purified to homogeneity from Xenopus oocytes (Engelke et al., 1980), and consists of a single polypeptide of approximately 38,500 daltons (Bieker et al., 1984). Under transcription conditions a single molecule of TFIIIA binds to the 50 bp long internal control region of the 5S genes (Bogenhagen et al., 1980; Bieker and Roeder, 1984; Smith et al., 1984). The resulting (metastable) complex is further stabilized by the sequential binding of fators IIIC and IIIB (Lassar et al., 1983; Bieker and Roeder, submitted). In addition, TFIIIA interacts with 5S RNA to form 7S ribonucleoprotein particles (7S RNPs) which serve as storage particles for 5S RNA during oogenesis (Honda and Roeder, 1980; Pelham and Brown, 1980). Therefore, TFIIIA is likely to be a key regulator of 5S RNA gene expression, both in its role as an initiation factor for 5S RNA gene transcription and, potentially, as a stabilizing component of 5S RNA in 7S RNPs. It has been suggested that the stable interaction of TFIIIA with 5S RNA in oocytes (and possibly in somatic cells) may provide an autoregulatory mechanism for 5S gene transcription (Honda and Roeder, 1980; Pelham and Brown, 1980).

As a result of the activation of both oocyte-type and somatic-type 5S genes, 5S RNA is maximally accumulated early in oogenesis. Late in oogenesis and early in embryogenesis little or no 5S RNA is synthesized (Ford, 1971). Although there is transient expression of both oocyte- and somatic-type 5S genes at mid-blastula, only somatic-type 5S RNA is synthesized in post-blastula stages and in adult tissues (Wormington et al., 1983). The accumulation of TFIIIA is also developmentally regulated (Pelham and Brown, 1980; Shastry et al., 1984). Early in oogenesis TFIIIA is present at approximately 10^{12} molecules per cell, thereby ensuring the continued activation of all 5S genes (approximately 100,000) in the face of accumulating 5S RNA and recruitment of TFIIIA into 7S RNPs. The cellular level of TFIIIA decreases late in oogenesis and then further from fertilization throughout embryogenesis. It has been suggested that the overall capacity of a cell for 5S RNA synthesis may be directly related to its level of TFIIIA. In particular, the selective transcription of somatic-type 5S genes in late embryos and adult somatic cells may be related to the low, potentially rate-limiting amount of TFIIIA in these cells.

Interestingly, late stage embryos and somatic cells contain a second protein, designated TFIIIA′, which is several-fold more abundant than TFIIIA in these cells. Although similar to TFIIIA by immunological criteria and CNBr peptide mapping, TFIIIA′ appears to be 2000 daltons larger by SDS polyacrylamide gel electrophoresis (Pelham et al., 1981; Roeder et al., 1981; Shastry et al., 1984). When partially purified, TFIIIA′ was not able to substitute for TFIIIA in cell-free transcription of 5S genes (Roeder et al., 1981; Shastry et al., 1984), but the possibility remains that it plays an important role in the selective expression of somatic-type genes or the general regulation of 5S gene expression in late embryonic and adult somatic cells.

This study was undertaken to investigate both the developmental regulation of TFIIIA and the mechanism(s) by which TFIIIA regulates 5S gene expression during development. To pursue these objectives, a cDNA clone of TFIIIA was isolated, characterized, and used in initial studies of the TFIIIA gene and in analyses of TFIIIA RNA levels throughout Xenopus development. In addition, characterization of this cDNA has allowed deduction, for the first time, of the complete amino acid sequence of a positively acting eucaryotic transcription factor and of a protein that specifically binds both DNA and RNA. Examination of the TFIIIA cDNA sequence revealed two potentially interesting homologies between the TFIIIA cDNA and RNA polymerase III transcripts/genes. The first involves a stretch of nucleotides that repeats twice near the 5′ end of the TFIIIA coding region and exhibits homology to the "B-box" of the Xenopus tmet gene promoter. The second homology is between

a stretch of nucleotides toward the 3′ end of the TFIIIA cDNA coding sequence and the Xenopus 5S gene internal control region. The implications of these homologies for possible mechanisms of TFIIIA and 5S gene regulation are discussed.

Results

Determination of Partial Protein Sequence of TFIIIA

TFIIIA is easily purified from immature ovaries in which it constitutes as much as 1%–10% of the total oocyte protein. Therefore the strategy for isolation of a cDNA clone of TFIIIA was to determine a partial amino acid sequence for the protein, synthesize sets of corresponding oligonucleotides, and use these to probe a Xenopus ovary cDNA library.

TFIIIA appears to be blocked at its amino-terminus (unpublished data). We therefore cleaved TFIIIA with cyanogen bromide (CNBr) and separated the resulting peptides by SDS polyacrylamide gel electrophoresis. Four peptides were identified—two of approximately 10,000 daltons, one of approximately 20,000 daltons, and a partial digestion product of approximately 30,000 daltons. Gas-phase microsequencing (Hunkapiller et al., 1983a) of these peptides was kindly performed by Drs. Michael Hunkapiller and Leroy Hood (California Institute of Technology). Only the 20 kd peptide and the 30 kd partial digestion product yielded sequence information. These two peptides produced the same amino acid sequences from their amino termini. Based on these results, we derived the CNBr peptide map pictured in Figure 1A. The determined amino acid sequence is presented in Figure 1B and its location within the entire TFIIIA protein is marked in Figure 1A.

Isolation of a cDNA Clone Encoding TFIIIA

Two regions within this sequence were chosen to direct synthesis of oligonucleotides for use as probes in cDNA library screening. Oligonucleotide set I (Oligo I) is a collection of 17 base long sequences corresponding to all possible nucleotide sequences derivable from amino acids 7–12, except for the third base of amino acid 12 (see Figure 1B). Oligonucleotide set II (Oligo II) is the entire set of 20-mers derivable from the sequence of amino acids 17–23, except for the third base of amino acid 23 (Figure 1B). These oligonucleotides were synthesized by standard phosphotriester chemistry (Myoshi et al., 1980), purified by G-50 Sepharose chromatography and sequenced for confirmation by the chemical cleavage method (Maxam and Gilbert, 1980).

Oligonucleotides I and II were used to screen a cDNA library, which was a kind gift from Dr. Douglas Melton (Harvard University). The library was constructed in the vector λgt10, from poly(A)$^+$ RNA of defolliculated oocytes of Xenopus laevis mature ovaries (personal communication, D. Melton). We screened approximately one million independent recombinants at high density, probing duplicate filters in parallel with Oligos I and II. Screening conditions (see Experimental Procedures) were determined by probing filters that contained a previously isolated Xenopus histone gene (M. Perry and R. Roeder, unpublished data) subcloned into λgt10, with a set of 17 base long oligonucleotides corresponding to a known region of the appropriate Xenopus histone sequence. This set of 17-mers had a nucleotide composition similar to that of the TFIIIA-derived oligonucleotides and was composed of an equivalent number of sequences (i.e., 48). It should be noted that the hybridization and wash temperatures employed were sig-

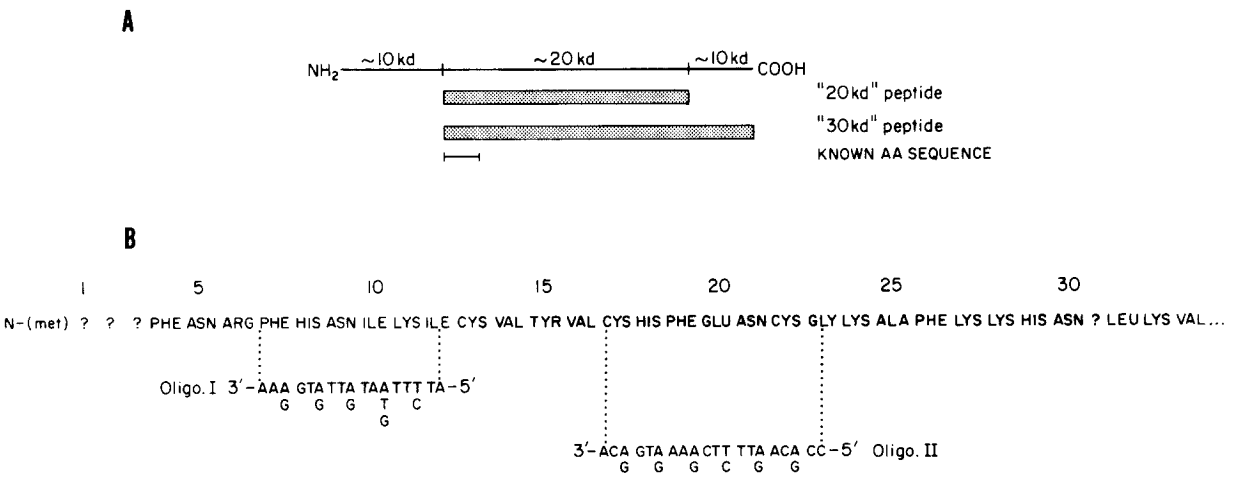

Figure 1. CNBr Peptide Map of TFIIIA and Partial Amino Acid Sequence as Determined by Peptide Sequencing

(A) CNBr peptide map of TFIIIA with derived positions of the 20 kd and 30 kd peptides marked. The position of the amino acid sequence determined by N-terminal microsequencing of the 20 kd and 30 kd peptides is denoted by '(⊢).

(B) N-terminal amino acid sequence of the 20 kd and 30 kd CNBr peptides and definitions of Oligos I and II. The nucleotide sequences of the families of oligonucleotides, Oligos I and II, were derived from the corresponding amino acids (whose boundaries are marked with vertical dotted lines). The identity of the first three amino acid residues could not be determined from the microsequencing data. (The N-terminal methionine is assumed because CNBr is known to cleave after methionines.)

nificantly lower than would have been predicted by the formula Tm_{app} = 4°(G:C) + 2°(A:T) (Suggs et al., 1981). This formula predicted a range of melting temperatures for the histone probe of 42°C–52°C. When temperatures within this range were used to screen the histone positive control system, we could not reliably detect positive signals. We determined experimentally that the optimal temperature was 33°C, and therefore used this temperature in screening the cDNA library with Oligos I and II.

Oligo I hybridized to approximately 250 plaques of the one million screened, and Oligo II to approximately 375. However, only 37 of the one million appeared to hybridize to both Oligos I and II. These 37 were picked and rescreened at lower density. After three rounds of screening, 11 independent clones were identified that hybridized relatively strongly and reproducibly to both probes in plaque hybridizations. Phage DNA was purified from these 11, digested with Eco RI, and analyzed by agarose gel electrophoresis and Southern blot hybridization using both Oligos I and II. All 11 contained at least one insert-derived Eco RI fragment that hybridized specifically to these two sets of oligonucleotides (data not shown). The clone containing the longest hybridizing insert fragment (λ3a1.b; approximately 1500 bp insert) was further characterized.

Characterization of cDNA clone λ3a1.b

A partial restriction map of λ3a1.b insert DNA is presented in Figure 2. To confirm that this clone contained DNA for TFIIIA, Oligo II was used as a primer for dideoxy sequencing of purified insert DNA. The sequence obtained corresponded exactly to the determined amino acid sequence. The λ3a1.b Eco RI insert was subcloned into pUC13. This subclone (designated "puc3a1.b") was the DNA source for the experiments described below. The complete nucleotide sequence of this cDNA insert was determined by the chemical cleavage method following the strategy depicted in Figure 2. The nucleotide sequence of the puc3a1.b insert DNA and the inferred amino acid sequence are presented in Figure 3. (Positions of "stop" codons in the other two reading frames are indicated.) The 27 amino acids determined by peptide sequencing correspond to amino acids 94–120 and nucleotides 280–360 of this sequence. The positions of methionines in the derived sequence are marked. They dictate a CNBr peptide map that agrees very closely with that determined experimentally for TFIIIA—both in relative position and in sizes of the peptides (an N-terminal peptide of approximately 9790 daltons, an internal peptide of approximately 19,250 daltons, and a C-terminal peptide of approximately 8800 daltons).

These data suggest that puc3a1.b contains the entire coding sequence of TFIIIA mRNA. At nucleotides 1059 to 1080 there are "stop" codons in all three reading frames, indicating the 3′ end of the coding sequence. These are followed by approximately 435 bp of untranslated sequence. This clone does not, however, contain a poly(A) stretch in the 3′ untranslated region of its noncoding strand. It also does not appear to include an AATAAA polyadenylation signal. Since these are both expected features of eucaryotic mRNA-derived cDNAs (Nevins, 1983), it is likely that puc3a1.b does not contain the entire 3′ untranslated sequence of the TFIIIA mRNA. These sequences were most likely lost during double-stranded cDNA synthesis. (First strand synthesis was primed with oligo-d(T) from stringently selected poly(A)$^+$ RNA template; D. Melton, personal communication.)

To determine whether the 5′ end of the cDNA clone contains the entire 5′ end of TFIIIA mRNA, a 15-mer complementary to nucleotides 74 to 88, "Oligo III," (see Figures 2 and 3) was used to prime synthesis from poly(A)$^+$

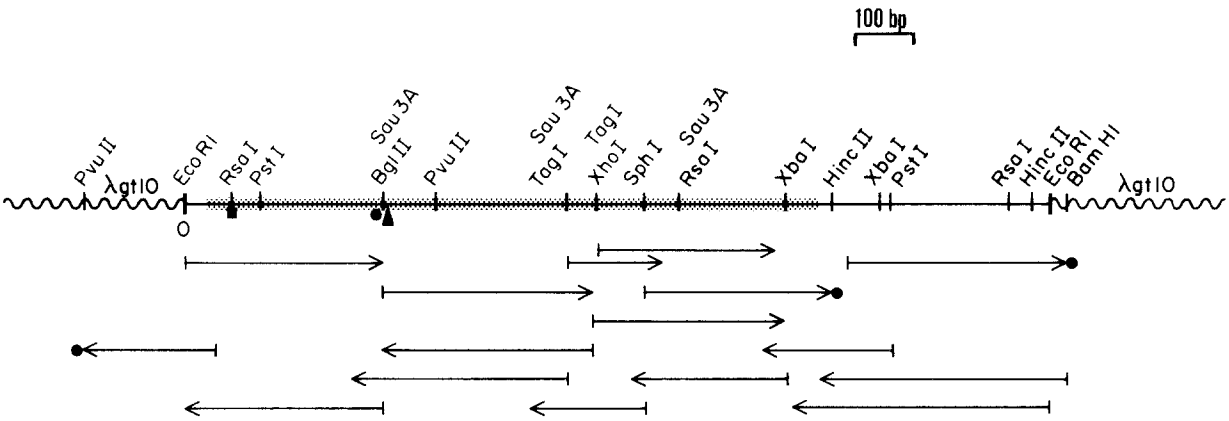

Figure 2. Partial Restriction Map of λ3a1.b DNA and Nucleotide Sequencing Strategy

The restriction map of this clone was determined by standard techniques (Maniatis et al., 1982). Arrows point in the 3′ to 5′ direction. Therefore arrows pointing to the right represent determination of coding strand sequences and arrows pointing to the left represent determination of noncoding strand sequences of this cDNA. (Strand-sense was originally assigned in relation to the known amino acid sequence for TFIIIA.) Lengths of arrows represent extent of readable nucleotide sequence resulting from a given sequencing reaction. All reactions represented in this figure were done according to Maxam and Gilbert (1980). Dots on arrowheads (⇀) signify 5′ end-labeling of DNA template. All other templates were 3′ end-labeled. The stippled area (▩) represents the proposed coding sequence of 3a1.b cDNA. The circle just below the restriction map (●) indicates the approximate position of Oligo I within the proposed coding sequence, the triangle (▲) indicates that of Oligo II, and the square (■) indicates that of Oligo III.

```
                                    10          20          30  *       40
                                GAATTCCGGA AGCCGAGGGC TGTTCAGTTG CTGAAGGAGA G

              50              60              70  *           80
   ATG GGA GAG AAG GCG CTG CCG GTG GTG TAT AAG CGG TAC ATC TGC TCT
 1 MET Gly Glu Lys Ala Leu Pro Val Val Tyr Lys Arg Tyr Ile Cys Ser    16
                                                3'|_____________|5'OligoIII
   90            100             110   *       120             130
   TTC GCC GAC TGC GGC GCT GCT TAT AAC AAG AAC TGG AAA CTG CAG GCG
17 Phe Ala Asp Cys Gly Ala Ala Tyr Asn Lys Asn Trp Lys Leu Gln Ala    32

     140             150             160             170       *   180
   CAT CTG TGC AAA CAC ACA GGA GAG AAA CCA TTT CCA TGT AAG GAA GAA
33 His Leu Cys Lys His Thr Gly Glu Lys Pro Phe Pro Cys Lys Glu Glu    48

        190*            200             210             220*            230
   GGA TGT GAG AAA GGC TTT ACC TCG CTT CAT CAC TTA ACC CGC CAC TCA
49 Gly Cys Glu Lys Gly Phe Thr Ser Leu His His Leu Thr Arg His Ser    64

           240             250             260       *   270             280
   CTC ACT CAT ACT GGC GAG AAA AAC TTC ACA TGT GAC TCG GAT GGA TGT
65 Leu Thr His Thr Gly Glu Lys Asn Phe Thr Cys Asp Ser Asp Gly Cys    80

   *     *   290             300             310*            320   *
   GAC TTG AGA TTT ACT ACA AAG GCA AAC ATG AAG AAG CAC TTT AAC AGA
81 Asp Leu Arg Phe Thr Thr Lys Ala Asn MET Lys Lys His Phe Asn Arg    96

   330     *   340             350             360             370
   TTC CAT AAC ATC AAG ATC TGC GTC TAT GTG TGC CAT TTT GAG AAC TGT
97 Phe His Asn Ile Lys Ile Cys Val Tyr Val Cys His Phe Glu Asn Cys   112
 3'|_____________|5' Oligo I              3'|_____________________
      380             390             400         * 410             420
   GGC AAA GCA TTC AAG AAA CAC AAT CAA TTA AAG GTT CAT CAG TTC AGT
113 Gly Lys Ala Phe Lys Lys His Asn Gln Leu Lys Val His Gln Phe Ser   128
   |5' Oligo II
          430             440             450             *460        * 470
   CAC ACA CAG CAG CTG CCA TAC GAA TGT CCT CAT GAA GGC TGT GAC AAG
129 His Thr Gln Gln Leu Pro Tyr Glu Cys Pro His Glu Gly Cys Asp Lys   144

           480             490       *   500       *   510             520
   CGG TTT TCT TTG CCT TCC CGT TTA AAA CGT CAT GAA AAA GTC CAT GCA
145 Arg Phe Ser Leu Pro Ser Arg Leu Lys Arg His Glu Lys Val His Ala   160

             530             540 *           550             560
   GGC TAT CCC TGC AAA AAG GAT GAT TCT TGC TCA TTT GTG GGA AAG ACT
161 Gly Tyr Pro Cys Lys Lys Asp Asp Ser Cys Ser Phe Val Gly Lys Thr   176

   570             580   *         590             600             610   *
   TGC ACA TTA TAC TTG AAA CAC GTG CCA GAA TGC CAT CAG GAC CTA GCA
177 Trp Thr Leu Tyr Leu Lys His Val Ala Glu Cys His Gln Asp Leu Ala   192

     620     *       630 *           640             650             660 *
   GTA TGT GAT GTG TGT AAT CGA AAA TTC AGG CAC AAA GAT TAC TTG AGG
193 Val Cys Asp Val Cys Asn Arg Lys Phe Arg His Lys Asp Tyr Leu Arg   208

         670             680             690             700             710
   GAT CAT CAG AAA ACT CAC GAA AAA GAG CGA ACT GTG TAT CTC TGC CCT
209 Asp His Gln Lys Thr His Glu Lys Glu Arg Thr Val Tyr Leu Cys Pro   224

           720             730             740             750       *   760
   CGA GAT GGC TGT GAC CGC TCC TAT ACC ACT GCA TTC AAT CTT AGA AGC
225 Arg Asp Gly Cys Asp Arg Ser Tyr Thr Thr Ala Phe Asn Leu Arg Ser   240

             770           * 780             790             800   *
   CAT ATA CAA TCA TTT CAT GAG GAA CAG AGA CCT TTT GTT TGT GAG CAT
241 His Ile Gln Ser Phe His Glu Glu Gln Arg Pro Phe Val Cys Glu His   256

   810             820             830       *     840       *   850
   GCT GGC TGC GGG AAA TGC TTT GCA ATG AAA AAA AGC CTA GAA AGA CAT
257 Ala Gly Cys Gly Lys Cys Phe Ala MET Lys Lys Ser Leu Glu Arg His   272

     860           * 870             880           *890             900
   TCA GTT GTA CAT GAT CCA GAG AAG AGG AAG CTG AAG GAG AAA TGC CCT
273 Ser Val Val His Asp Pro Glu Lys Arg Lys Leu Lys Glu Lys Cys Pro   288

         910             920             930             940             950
   CGC CCA AAG AGA AGC CTG GCC TCT CGC CTC ACT GGA TAC ATA CCC CGC
289 Arg Pro Lys Arg Ser Leu Ala Ser Arg Leu Thr Gly Tyr Ile Pro Pro   304

           960             970             980             990           *1000
   AAG AGC AAA GAA AAA AAT GCA TCC GTT TCG GGA ACA GAA AAG ACT GAT
305 Lys Ser Lys Glu Lys Asn Ala Ser Val Ser Gly Thr Glu Lys Thr Asp   320

            *1010      *   1020          1030*          1040
   TCA CTT GTG AAA AAT AAG CCC TCT GGC ACT GAA ACA AAT GGC TCA TTG
321 Ser Leu Val Lys Asn Lys Pro Ser Gly Thr Glu Thr Asn Gly Ser Leu   336

   1050  *     *1060 *         1070    *   *1080         1090*       1100
   GTT CTA GAT AAA TTA ACT ATA CAA TAATATA AGAAAACATT TAAATTTATT
337 Val Leu Asp Lys Leu Thr Ile Gln  *

            1110*       1120        1130 *      1140  *     1150        1160
   TTTTTATTTG TTAAAATTGC CCTCAGGATG GTTAACCCAT ATTTAGTGTG GGTTTTTCT

            1170  *     1180        1190    *   1200        1210  *     1220
   TTTTTTACAG CTTTAATTCA TTTTTTTTCG GCTATAACAA AAGGAATCTG TTCTAGACGC

     *       1230  *     1240        1250        1260 *      1270      * 1280
   ATGATTTGTT TTATGAACTG CAGTATTGGC CATGCCTACA GGTAAAGGCA CAGTGTTAAT
                                                   *
            1290        1300        1310  *     1320        1330        1340
   GGCTACATAC CTCTTCTACC CCATGTTTGC TATTAAAAGT GAGGTGCAGC AGCCACTGGT

            1350        1360    *   *1370       1380        1390        1400
   CTGTTTATTT ACAATACATT CATTTAGTAA GACTCTGTAT TCATTTTCAA AAGAATCACT

   *        1410        1420        1430  *     1440        1450        1460
   AAGGGAATGT GCAAAATTGT TATCACTCTA CTGTAAACAC AAATGTACTG CTTGCACCCT

            1470        1480      * 1490*       1500        *1510
   GTTGGTGGGG CTTTTTTTGG GGAGGTTGAC TGACCCTGTT TTTTTTTTAA CGGAATTC
```

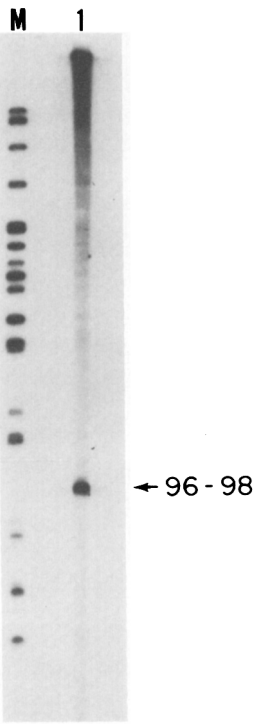

Figure 4. Comparison of the 5′ Ends of puc3a1.b cDNA and Its Corresponding mRNA by Primer-Extension Analysis

Lane 2 shows the electrophoretic separation of DNA synthesized by priming with Oligo III on a Xenopus oocyte poly(A)$^+$ RNA template. The major product, indicated with an arrow, is 96–98 bases long as determined by comparison with size markers (lane 1, pBR322 digested with Hpa II).

RNA template to the 5′ end of the RNA. The template was total poly(A)$^+$ RNA prepared from previtellogenic oocytes. If the 5′ end of the cDNA were identical to that of the corresponding RNA, the product length should have been 88 bases. Figure 4 indicates that the actual product was 96–98 nucleotides long. S1 nuclease mapping (data not shown) demonstrated that all but approximately ten bases of this primer-extension product were protected from S1 nuclease digestion by the 5′ terminal Pst I restriction fragment of the puc3a1.b insert DNA (see Figure 2). These findings imply that puc3a1.b is missing only approximately ten bases of 5′ transcribed but untranslated sequence, that it contains the complete coding sequence for TFIIIA, and that nucleotides 42 to 44 (Figure 3) represent the initiating ATG. (The formal possibility remains, however, that the missing ten nucleotides at the 5′ end of this clone contain another ATG at which translation might initiate.) Table 1 compares the amino acid compositions of the

Figure 3. Nucleotide Sequence of puc3a1.b Eco RI Insert DNA and Derived Amino Acid Sequence of TFIIIA

The nucleotide sequence given corresponds to the noncoding strand of the cloned cDNA. Nucleotide #1 represents the first base in the pUC13 Eco RI cloning site. Nucleotides #42–44 correspond to the codon for the proposed initiating methionine. Asterisks (*) mark the positions of all "stop" codons in all three reading frames. Underlined nucleotides (in the 3′ transcribed but untranslated region) denote bases that may be sequencing artifacts. The sequences that would hybridize to Oligos I, II, and III are noted.

protein predicted to be encoded by puc3a1.b and of oocyte 7S RNP-derived TFIIIA as determined by Picard and Wegnez (1979). These values agree to within 20% for all amino acids and to within 10% for 12 of 20. The fact that puc3a1.b encodes a 27 amino acid stretch identical to a known amino acid sequence of TFIIIA (Figures 1B and 3), the agreement between the theoretical CNBr map of the puc3a1.b gene product and the experimentally derived CNBr map of TFIIIA (Figures 1A and 3), and the close similarity between the cDNA-derived and the experimentally derived amino acid compositions (Table 1) establish that puc3a1.b cDNA encodes the Xenopus transcription factor TFIIIA.

Analysis of TFIIIA Sequences in Xenopus laevis Genomic DNA

To estimate the number of copies of the TFIIIA gene in the Xenopus genome, Southern blot hybridization analyses were performed on Xenopus erythrocyte high molecular weight DNA. DNA was digested to completion with each of several restriction enzymes that do not cut within the TFIIIA cDNA sequence and then probed with nick-translated puc3a1.b Eco RI insert DNA. A relatively simple hybridization pattern, consisting of three to four bands, was observed in each analysis (Figure 5A). Quantitation of the hybridization signals (see Experimental Procedures and legend to Figure 5) indicated that the gene encoding TFIIIA is present at one (or a very few) copies per haploid genome. Figure 5B contains an autoradiogram of a Southern blot analysis of a double digest of Xenopus genomic DNA (lane g) and puc3a1.b insert DNA (left three lanes) with enzymes (Pst I and Pvu II) that do cut within the TFIIIA cDNA sequence. Note that none of the five bonds visible in the genomic DNA (indicated by arrows, with their lengths in base pairs) migrates with the 300 bp or the 785 bp long internal fragments of the cDNA insert (Figure 5B, lanes 1, 3, and 6). These results suggest that the TFIIIA gene contains introns within both the 300 bp and 785 bp Pst I/

Table 1. Codon Usage and Comparison of Amino Acid Compositions

	Codon[a] Usage		puc3a1.b[a] Amino Acid Composition	TFIIIA[b] Amino Acid Composition		Codon[a] Usage		puc3a1.b[a] Amino Acid Composition	TFIIIA[b] Amino Acid Composition
Ala	GCT	3	15	18	Phe	TTT	10	17	18
	GCC	2				TTC	7		
	GCA	8			Leu	TTA	5		
	GCG	2				TTG	5		
Tyr	TAT	6	11	10.5		CTT	3	25	25
	TAC	5				CTC	3		
His	CAT	16	25	24		CTA	3		
	CAC	9				CTG	6		
Lys	AAA	23	43	41	Ile	ATT	0		
	AAG	20				ATC	3	6	4
Gln	CAA	3				ATA	3		
	CAG	7			Met	ATG	3	3	2.5
			32	39	Val	GTT	5		
Glu	GAA	12				GTC	2	17	13
	GAG	10				GTA	2		
Asn	AAT	6				GTG	8		
	AAC	7			Ser	TCT	5		
			29	32		TCC	3		
Asp	GAT	10				TCA	6	22	22
	GAC	6				TCG	3		
Cys	TGT	11	23	15		AGT	1		
	TGC	12				AGC	4		
Trp	TGG	2	2	2	Pro	CCT	5		
Arg	CGT	2				CCC	4	15	21
	CGC	4				CCA	5		
	CGA	3	20	20		CCG	1		
	CGG	2			Thr	ACT	11		
	AGA	6				ACC	3	21	18
	AGG	3				ACA	7		
Gly	GGT	0				ACG	0		
	GGC	10	18	19					
	GGA	7							
	GGG	1							

[a] For protein derived from puc3a1.b cDNA sequence, as depicted in Figure 3.

[b] Amino acid composition reported by Picard and Wegnez, 1979 for oocyte 7S RNP-derived TFIIIA; corrected to total of 344 amino acids.

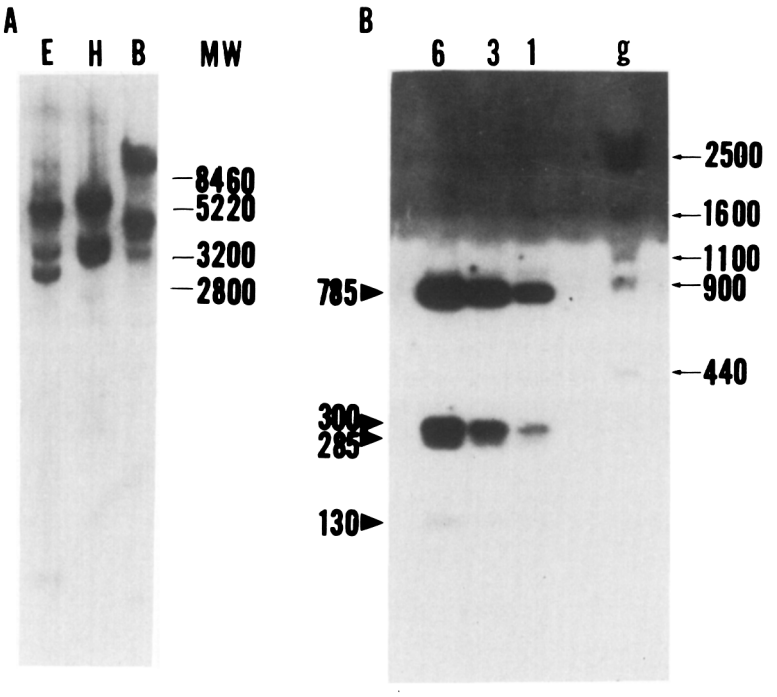

Figure 5. Analyses of TFIIIA Sequences in Xenopus Genomic DNA

Southern blot hybridization analyses of Xenopus erythrocyte high molecular weight DNA.

(A) DNA was digested to completion with Eco RI, lane E, Hind III, lane H, and Bgl I, lane B.

(B) DNA was digested with Pst I and Pvu II. Lane g, Xenopus erythrocyte genomic DNA. Lanes 1, 3, and 6 contain 1, 3, and 6 haploid genome equivalents of puc3a1.b Eco RI insert DNA, respectively. Lanes E, H, B, and g contain approximately 20 μg DNA each. Following digestion, all DNAs were separated electrophoretically, transferred to nitrocellulose filters and probed with nick-translated puc3a1.b Eco RI insert DNA. Signals of autoradiograms in (A) and (B) were quantitated by densitometric scanning of the resulting autoradiograms and comparison to puc3a1.b insert DNA standards present on each filter (calculations of peak areas were based on valley-to-valley measurements). For example, the relative signal intensities (sum of all bands in each lane) for the autoradiogram in (B) are: lane 1 = 1.0, lane 3 = 4.1, lane 6 = 6.3, lane g = 1.3. The dark background at the top of the filter in (B) is an artifact of hybridization which was unique to this filter.

Pvu II regions of the coding sequence. This experiment also confirmed the low copy number of the TFIIIA sequence (see legend to Figure 5).

TFIIIA RNA Levels in Oogenesis and Embryogenesis

As an initial approach to studying the regulation of TFIIIA gene expression in Xenopus laevis development, the steady-state levels of TFIIIA RNA in staged oocytes and embryos were analyzed by Northern blot hybridizations. Total RNA was prepared from each stage in the presence of guanidinium isothiocyanate and RNA equivalents of a fixed number of oocytes or embryos were loaded on formaldehyde-agarose gels and separated electrophoretically. RNAs were then transferred to nitrocellulose and filters treated under stringent conditions (see Experimental Procedures). Nick-translated puc3a1.b insert DNA was used to probe the filters. A major band of approximately 1550 nucleotides was detected in each stage of oogenesis and in early embryos (Figures 6A–6D). Note that the length of this RNA is in close agreement with the size of the cDNA insert in puc3a1.b and suggests that only a relatively small amount (<50 bases) of 5′ and 3′-transcribed but untranslated sequences are missing from the clone. Even after longer exposures of the autoradiograms no signal was detectable in neurulae or adult liver. (See below and Discussion for comments on the size and level of the RNA in swimming tadpoles and Xenopus adult kidney tissue culture cells.)

Levels of RNA were quantitated by densitometric scanning of autoradiograms and are summarized in Table 2. On a per oocyte/embryo basis, TFIIIA RNA appeared to be most abundant in Stage I, II, and III oocytes (Figures 6A and 6B). Between stages III and IV the level of TFIIIA RNA decreased about 5-fold (Figure 6B). The abundance of TFIIIA RNA was then approximately constant, at this reduced level, through stages IV–VI (Figure 6B). Moreover, the abundance of TFIIIA RNA in fertilized eggs and blastulae appeared to be equivalent to that of Stage IV–VI oocytes on a per embryo versus per oocyte basis (Table 2). Since embryos do not become transcriptionally active until mid-blastula (Newport and Kirschner, 1982), it is likely that the TFIIIA RNA from late stage oocytes was retained through oocyte maturation, fertilization, and early cleavage. Between blastula and neurula formation, the per embryo level of TFIIIA RNA decreased dramatically (to undetectable levels; Figure 6C), presumably as a result of mRNA degradation. Similar analyses of both swimming tadpole and a Xenopus kidney cell line (Figures 6C and 6D) indicated that these cells contain low, but detectable, levels of TFIIIA RNA, suggesting that new synthesis occurred between neurulation and development into swimming tadpoles (Figure 6C). The possibility of new synthesis in late embryo and adult somatic cells is further supported by the apparently heterogeneous nature and larger size of the hybridizing RNA (Figures 6C and 6D, and Discussion). However, direct analysis of newly synthesized RNAs is required to resolve this question.

Discussion

Using short oligonucleotides as probes, we have isolated a cDNA clone which appears to contain the entire coding sequence for the Xenopus transcription factor TFIIIA, as well as most, but not all, of its 5′ and 3′ transcribed but untranslated regions. Nucleotide sequence analysis of this

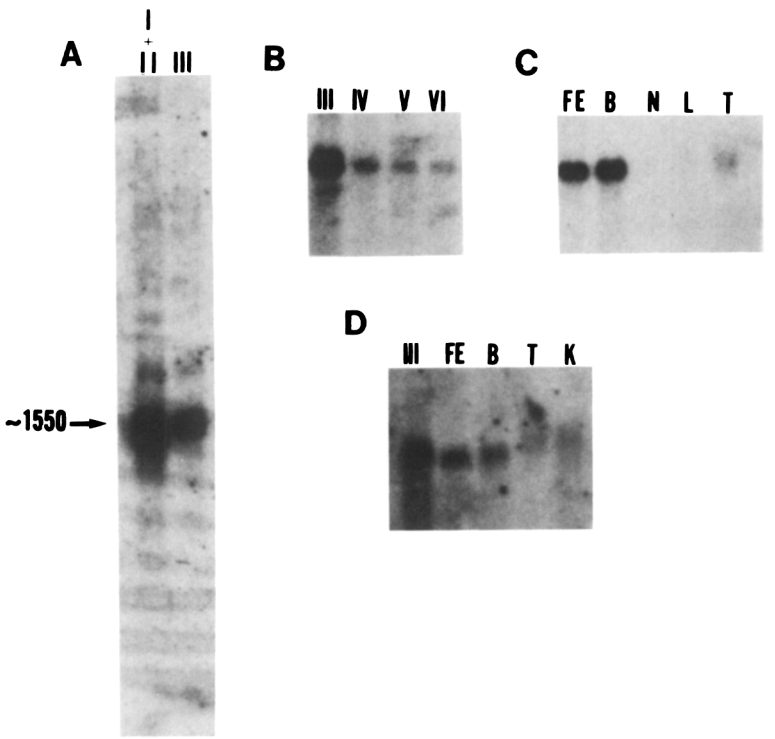

Figure 6. TFIIIA RNA Levels throughout Xenopus Development

Total cellular RNA was isolated from staged oocytes and embryos, from adult liver, and from actively growing cells of an adult kidney cell line. RNA was analyzed by Northern blot hybridizations using nick-translated puc3a1.b Eco RI insert DNA as probe.

(A) Lanes I and II, Stage I and II oocytes; lane III, Stage III oocytes. Forty oocyte equivalents were loaded in each lane. Autoradiographic exposure was for 12 hr on pre-flashed XAR film, with an intensifying screen, at −80°C.

(B) Lane III, Stage III oocytes; lane IV, Stage IV oocytes, lane V, Stage V oocytes; lane VI, Stage VI oocytes. Twenty oocyte equivalents were loaded in each lane. Autoradiographic exposure was for 15 hr on XAR film, with an intensifying screen, at −80°C.

(C) Lane FE, fertilized eggs; lane B, blastulae; lane N, neurulae; lane L, adult liver; lane T, swimming tadpoles. Forty embryo equivalents were loaded in the appropriate lanes and 100 μg of adult liver RNA in lane L. Autoradiographic exposure was as for (B).

(D) Lane III, Stage III oocytes; lane FE, fertilized eggs; lane B, blastulae; lane T, swimming tadpoles; lane K, adult kidney cell line. Fifteen oocyte or embryo equivalents were loaded in lanes III, FE, and B. Fifty embryo equivalents were loaded in lane T, and 20 μg of kidney cell RNA was loaded in lane K. Note that lanes III, FE, and B contained approximately equal amounts of total RNA. Autoradiographic exposure was for 6½ days on XAR film, with an intensifying screen, at −80°C. 28S and 18S RNAs served as size markers for these analyses. Signals were quantitated by densitometric scanning of autoradiograms and comparison to known amounts of identically treated puc3a1.b insert DNA. Calculations of peak areas were based on valley-to-valley measurements.

clone has enabled us to deduce the complete amino acid sequence of TFIIIA. This represents the first time that the entire primary structure of a positively acting eucaryotic transcription factor has been deduced. TFIIIA is also the only protein of which we are aware that binds to specific sites on both DNA and RNA in vivo (Honda and Roeder, 1980; Pelham and Brown, 1980; Pieler and Erdmann, 1983). Its primary structure is therefore of particular interest. Preliminary sequence comparisons have failed to demonstrate significant nucleotide or amino acid sequence homology between TFIIIA and other DNA binding proteins (λ*cro* and λ*cI*, E. coli lac repressor, and E. coli CAP).

In view of the dramatic difference in the level of TFIIIA in immature oocytes (high) versus somatic cells (low), and because of the interaction of TFIIIA with the multicopy, two-type system of 5S RNA genes (20,000 oocyte-type and 400 somatic-type 5S RNA genes per haploid genome), it was of interest to determine the number of copies of the TFIIIA sequence in the Xenopus genome. Xenopus genomic DNA was therefore analyzed by quantitative Southern blot hybridizations, which indicated that the gene encoding TFIIIA is present at one (or a small number of) copies per haploid genome. Thus it is likely that one (or a few) genes for TFIIIA are transcribed at a relatively high rate in early oocytes and less actively in somatic cells. These analyses, although suggesting that there is only one TFIIIA gene, do not rule out the possibility that there are distinct oocyte-type and somatic-type genes encoding TFIIIA. The blot hybridizations were performed under sufficiently stringent conditions such that the puc3a1.b probe might not have hybridized to related but only partially homologous genes. Furthermore, spectrophotometric quantitation of autoradiographic signals is not precise enough to distinguish one from two gene copies with absolute reliability. Moreover, comparison of DNA content and chromosome number among various Xenopus species has led to the suggestion that a chromosome duplication occurred relatively early in Xenopus evolution and that Xenopus laevis is therefore essentially tetraploid (Bisbee et al., 1977). If this is true, virtually all genes in Xenopus laevis should be present in at least two copies per haploid genome. A Xenopus genomic library is currently being probed with puc3a1.b cDNA in order to characterize and quantitate further the TFIIIA gene(s).

Analyses of TFIIIA RNA levels throughout development indicate that the pattern of expression observed for TFIIIA RNA correlates well with that described for TFIIIA protein (Wormington et al., 1983; Shastry et al., 1984) and 5S RNA (reviewed in Korn, 1982). Early oocytes accumulate very high levels of 5S RNA (10^{11} to 10^{12} molecules per cell). To do so they presumably require high levels of TFIIIA protein (which is, in fact, also present at 10^{11} to 10^{12} molecules per cell). Significantly, it appears to be young oocytes that most actively synthesize and accumulate TFIIIA RNA. However, whereas TFIIIA represents 1%–10% of the protein in young oocytes, it appears to comprise only 0.01%–0.1%

Table 2. Levels of TFIIIA RNA and Protein throughout Xenopus Development

Oocyte Stage[a]	Approximate Cell Number	Molecules of TFIIIA RNA[b] per		Molecules of TFIIIA Protein[c] per	
		Oocyte/Embryo	Cell	Oocyte/Embryo	Cell
I + II	1	5×10^6	5×10^6	$8\text{–}10 \times 10^{11}$	$8\text{–}10 \times 10^{11}$
III	1	5×10^6	5×10^6	9×10^{11}	9×10^{11}
IV	1	1×10^6	1×10^6	7×10^{11}	7×10^{11}
V	1	1×10^6	1×10^6	3×10^{11}	3×10^{11}
VI	1	1×10^6	1×10^6	4.5×10^{10}	4.5×10^{10}
Embryo Stage[a]					
Fertilized egg	1	1×10^6	2×10^6	3×10^9	3×10^9
Cleavage	8	–	–	4.4×10^9	5.5×10^8
Blastula	4,000	2×10^6	5×10^2	–	–
Gastrula	50,000	–	–	4.5×10^9	9×10^4
Swimming tadpole	500,000	3×10^5	≤ 1	1.5×10^9	3×10^3

[a] For a given set of experiments, staged oocyte RNAs were isolated from a single ovary. Similarly, staged embryo RNAs were isolated from embryos resulting from a single mating of a second frog. Absolute numbers of RNA molecules varied from experiment to experiment, but the same pattern was observed in each case. The numbers presented here are from a representative set of experiments.
[b] Levels of TFIIIA RNA were determined by quantitation of Northern blot hybridization results (Figure 6) as described in Experimental Procedures.
[c] Levels of TFIIIA protein are quoted from Shastry et al., 1984.

(weight:weight) of the hnRNA plus mRNA of these cells (assuming hnRNa plus mRNA is approximately 1% of the total RNA in early oocytes; Rosbash and Ford, 1974). The data presented indicate that TFIIIA protein in early oocytes is derived from approximately five million copies of TFIIIA RNA. The abundance of this RNA decreases about 5-fold in Stage IV oocytes and is apparently retained at this level (although presumably in a nonfunctional state) through fertilization and early cleavage. After blastula formation the abundance of TFIIIA RNA is severely reduced, presumably as a result of degradation. The level of TFIIIA protein also decreases during late embryogenesis and may become limiting for 5S gene expression. Appropriate gene transfer experiments with the TFIIIA cDNA clone should make it possible to ascertain whether a limiting level of TFIIIA protein is responsible for the apparent selective repression of oocyte-type 5S RNA genes observed in somatic cells.

Swimming tadpoles and Xenopus adult kidney cells contain a species of puc3a1.b insert-hybridizing RNA that appears both more heterogeneous and longer than that found in oocytes and early embryos. This finding has several interesting implications. First, as discussed above, it indicates that new synthesis of TFIIIA-related RNA occurs both subsequent to neurulation (during development) and in adult tissues. Second, it suggests that variable degrees of polyadenylation or the presence of additional sequences within other areas of the RNA's noncoding or coding regions (e.g., the result of alternate processing) could distingiush the TFIIIA-related RNA in these cells from that of less mature cells. Third, the larger size of this RNA relative to that in oocytes and early embryos may correlate with the apparent size difference between TFIIIA and TFIIIA′ (50–150 nucleotides contain enough coding capacity for the 2000 daltons in question). The in vivo function of TFIIIA′ is unknown, but, among other possibilities (Pelham et al., 1981; Roeder et al., 1981; Shastry et al., 1984), it could be responsible in part for the selective expression of somatic-type 5S RNA genes in somatic cells. It is not clear whether TFIIIA′, TFIIIA, or both of these proteins are encoded by the larger TFIIIA-related RNA observed in mature embryos and adult somatic cells. We cannot rule out the possibility, for example, that these cells contain an undetectably low level of the 1550 nucleotide RNA in addition to the larger species. Furthermore, the apparent increased molecular weight of TFIIIA′ might be due to chemical modification, rather than to additional amino acids. Experiments to further characterize both this larger TFIIIA-related RNA and TFIIIA′ are underway in our laboratory. (Note that Stage I and II oocytes contain, in addition to the major signal at 1550 nucleotides, a faint, hybridizing band at approximately 1750 nucleotides [Figure 6A] which appears different in size from the tadpole and kidney cell hybridizing RNA. The biological signifiance, if any, of this 1750 nucleotide RNA in young oocytes is unclear, but it, too, could derive from a second, related gene or from differential processing of the TFIIIA transcript.)

These investigations of the developmental regulation of TFIIIA suggest that on a per oocyte/embryo basis, TFIIIA is maximally expressed in early oogenesis, after which the abundance of its RNA decreases several-fold to a level that remains constant until blastulation. Golden et al. (1980) have studied the pattern of accumulation of 500 individual poly(A)$^+$ RNAs during Xenopus oogenesis. All of the nonmitochondrial RNAs examined showed a pattern of maximal accumulation in early oogenesis and then a maintenance of this maximal level through oocyte maturation.

The clones analyzed did not show the several-fold decrease we observed for TFIIIA RNA between Stages III and IV. This difference may be due to the use of poly(A)$^+$ RNA in contrast to the use in this study of total cellular RNA, or, as reflected in the apparently greater abundance of the RNAs studied relative to that of TFIIIA, these RNAs may belong to a different (more prevalent) regulatory class.

Examination of the nucleotide sequence of the TFIIIA cDNA clone uncovered two homologies that may play a role in the developmental regulation of TFIIIA gene expression. First, similar stretches of nucleotides occur between positions 45–54 (GGAGAGAA) and 156–166 (GGAGAGAAACC) of TFIIIA cDNA. This sequence is potentially interesting not only because it repeats twice near the 5′ end of the TFIIIA coding region, but also because it differs in only two bases from the section of the internal control region of the Xenopus tmet gene known as the "B-box" (GGATCGAAACC) (Hofstetter et al., 1981). The significance of these six out of eight and nine out of 11 bp homologies is unclear, particularly given the very purine-rich nature of this short sequence (and the fact that these sequences demonstrate a lower degree of homology to the B-box consensus sequence; Galli et al., 1981). However, this homology might, for example, reflect temporally coordinated regulation of TFIIIA and pol III transcripts via binding of a common factor.

A second region of the TFIIIA coding sequences (nucleotides 750–793) demonstrates marked homology with the internal control region of the 5S gene (noncoding strand; Figure 7). Nucleotides 750–764 contain 13 of 15 bases that are identical to the 5′ border of the internal control region of Xenopus laevis oocyte 5S DNA, and nucleotides 782–793 contain nine of 12 bases identical to the 3′ border of this region. This region of homology suggests that TFIIIA could interact with its own gene or the derived RNA to autoregulate expression at the transcriptional or translational levels. Alternatively, these homologous sequences might enable another factor (e.g., TFIIIB, TFIIIC, or an as yet unidentified factor) to interact with both 5S and TFIIIA sequences to regulate their expression coordinately. These hypotheses are being tested using a combination of in vitro and in vivo transcription and translation systems and the cloned 5S and TFIIIA genes.

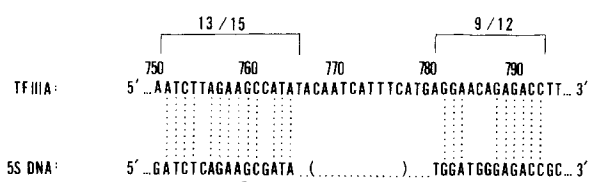

Figure 7. Comparison of TFIIIA cDNA and Xenopus 5S RNA Gene Sequences

The regions of sequence homology are depicted. The sequence of the noncoding strands of TFIIIA and 5S DNA are given. Nucleotide numbers refer to positions in Figure 3. The 5S gene sequence presented is that of Xenopus laevis oocyte 5S DNA (Miller et al., 1978).

Experimental Procedures

Preparation of CNBr Peptides and Peptide Sequencing

TFIIIA was prepared from Xenopus oocyte 7S RNPs (Dignam et al., 1983) and CNBr cleavage performed as in Shastry et al. (1984) for 22 hr at room temperature, in the dark. CNBr peptides of TFIIIA proved to be extraordinarily sticky and therefore yields of these peptides were low (1%–10%). The peptides were purified by SDS gel electrophoresis (Hunkapiller et al., 1983b) using the specified reagents and electroelution chamber. Microsequencing (Hunkapiller et al., 1983a) was performed by Dr. M. Hunkapiller on approximately 0.5 nmole of each peptide.

Synthesis of Oligodeoxynucleotides

Oligonucleotides I and II and the histone-derived probe were synthesized by the solid-phase phosphotriester method on a manual Bachem DNA synthesizer, (Myoshi et al., 1980). They were purified by G-50 Sepharose chromatography on a 2 cm × 50 cm column in 0.01 M triethylammonium bicarbonate (pH 7.5). One milliliter fractions were collected across the first peak. To assess purity a sample of each fraction was 5′ end-labeled with γ-^{32}P-ATP (Sgaramella and Khorana, 1972) and separated electrophoretically on a 12% polyacrylamide/7 M urea gel in TBE (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA).

Oligonucleotide III was synthesized on a Beckman automatic synthesizer, by Dr. Peter Model and colleagues (The Rockefeller University). This oligonucleotide was purified on a Mono Q FPLC column (Pharmacia). The column was pre-equilibrated in 12 mM NaOH (pH 11.5), 300 μg of Oligo III were loaded in 2.5 ml, and the size classes of oligonucleotides separated on a 0.1–1.0 M NaCl/12 mM NaOH (pH 11.5) gradient. Fractions were 5′ end-labeled and analyzed as above. Only fractions containing the longest oligonucleotides (15, 17, or 20 bases, as appropriate) were employed in subsequent experiments.

Screening of a cDNA Library with Oligonucleotide Probes

A cDNA library prepared from poly(A)$^+$ RNA of defolliculated oocytes of mature Xenopus ovaries in the vector, λgt10, was a generous gift from Dr. D. Melton, Harvard University. This library was screened using several modifications of standard procedures (Maniatis et al., 1982). One million recombinants were plated on C600hfl cells at high density—approximately 40,000 plaques per 150 × 25 mm plates. Two nitrocellulose (Schleicher and Schuell, BA85) replica filters were made per plate, the DNA was denatured and fixed to the filters, and the filters were baked at 80°C for 2 hr. After baking, filters were briefly wet in 6X SSC (1X SSC = 0.15 M NaCl, 0.015 M Na-citrate [pH 7.0]), prehybridized for 4 hr at 33°C in 6X SSC, 10X Denhardt's (1X Denhardt's = 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% BSA), 0.2 mg/ml sonicated salmon sperm DNA, and hybridized for 16–18 hr at 33°C in fresh 6X SSC, 10× Denhardt's, 0.2 mg/ml sonicated salmon sperm DNA (Sigma) containing 2.0 pmol/ml of kinased Oligo I or Oligo II (specific activity of probes = 1–3 × 10^8 Cerenkov cpm per pmol). After hybridization, filters were washed twice for 10 min at 4°C, then once for 10 min at room temperature, and finally once for 5 min at 33°C. 6X SSC was used for all washes. Filters were dried briefly at room temperature, enclosed in Saran Wrap, marked with radioactive ink and exposed for 15 hr to pre-flashed film in the presence of an intensifying screen, at −70°C. Plaques that appeared to hybridize to both Oligos I and II were picked and rescreened (through two more rounds) under identical conditions, but at lower density (50–1000 plaques per 100 mm plates).

Southern Hybridization Analyses of λgt10-cDNA Clones

Rapid lysate DNA was prepared (Davis et al., 1980) from each of the 11 cDNA clones that hybridized reproducibly to both Oligos I and II. A portion of each of these DNAs were digested to completion with Eco RI (a kind gift from Drs. W. Jack and P. Modrich) (Maniatis et al., 1982). The digested DNAs were separately electrophoretically on 1% agarose/TBE gels and blotted to nitrocellulose (Maniatis et al., 1982). These filters were probed independently with both Oligo I and Oligo II under the same conditions used for screening the cDNA library above.

Primer-Extended Dideoxy Sequence Analysis of λ3a1.b Insert DNA

Approximately 200–400 ng of purified (double-stranded) Eco RI insert DNA

was lyophilized with 1.2 μg of Oligo II (10:1 molar ratio, Oligo II: insert DNA, taking into account 64-fold degeneracy of Oligo II). The DNAs were resuspended in 4 μl of 1X Klenow buffer (Maniatis et al., 1982) and transferred to a glass capillary tube. The sealed tube was incubated at 100°C for 5 min, and then plunged immediately into ice water. Sequencing was performed (Sanger et al., 1977), and the reaction products analyzed on 8% polyacrylamide/7 M urea "thin" gels (Maxam and Gilbert, 1980).

Subcloning, Restriction Mapping, and DNA Sequencing of λ3a1.b (puc3a1.b) Insert DNA

Subcloning of insert DNA into pUC13 and restriction mapping were done according to Maniatis et al. (1982) using enzymes supplied by BRL and New England Biolabs. Sequence analysis of puc3a1.b insert DNA was done using chemical cleavage methodology (Maxam and Gilbert, 1983), following the strategy depicted in Figure 2. Radioactive reagents for all experiments were purchased from New England Nuclear.

Primer Extension of Oligo III from Poly(A)$^+$ RNA Template and S1 Nuclease Protection Analysis of the Product

Oligo III was labeled with γ-^{32}P-ATP using T4 kinase according to the supplier's instructions, to a specific activity of 2.8×10^8 Cerenkov cpm per μg. Kinased oligonucleotide was separated from unincorporated γ-^{32}P-ATP by G-50-40 Sepharose chromatography in the presence of 10 mM Tris-Cl (pH 7.5)/1 mM EDTA.

Primer-extended synthesis of cDNA from a poly(A)$^+$ RNA template was performed according to Agarwal et al. (1981). RNA was purified from previtellogenic ovaries of Xenopus laevis by the guanidinium/hot phenol method and poly(A)$^+$ RNA selected by one passage over oligo-d(T) cellulose (both procedures as described in Maniatis et al., 1982). Poly(A)$^+$ RNA was stored in ethanol at −80°C. For primer extension, 10 μg of poly(A)$^+$ RNA was mixed with 1 μg of kinased Oligo III (10:1 molar ratio, Oligo III:RNA). Preparative isolation of the primer-extension product was by electrophoresis on a 12% polyacrylamide/7 M urea slab gel (40 × 20 × 0.15 cm), and electroelution of the product onto DEAE membrane (Schleicher and Schuell, NA45) according to manufacturer's directions.

The primer-extension product was analyzed by S1 nuclease digestion. The 5′ terminal Pst I fragment of puc3a1.b insert DNA was purified from an agarose/TBE gel onto DEAE membrane (see above and Maniatis et al., 1982), denatured by heating at 90°C 15 min, quick-cooled in ice water, made 6X in SSC, and added to purified primer-extension product. Nine micrograms of carrier tRNA was added to the reaction (20 μl final volume), which was then incubated for 12 hr at 68°C. Following hybridization, S1 nuclease buffer (130 μl; 30 mM Na-acetate (pH 4.6), 50 mM NaCl, 1 mM ZnSO4, 5% glycerol), 20 μg carrier tRNA and 100 U S1 nuclease (BRL) were added, and the reaction was incubated at 37°C for 60 min. The reaction was stopped by the addition of EDTA and carrier tRNA, and the products ethanol-precipitated and analyzed on an 8% polyacrylamide/7 M urea "thin" gel (Maxam and Gilbert, 1980) (^{32}P-labeled Hpa II digested pBR322 fragments served as size markers.)

Southern Hybridization Analysis of Genomic DNA

Xenopus high molecular weight DNA purified from erythrocytes was a kind gift from Barkur S. Shastry. DNA was digested with 1 U of restriction enzyme per μg DNA, under supplier's recommended conditions, for at least 8 hr at 37°C. The extent of digestion was monitored by ethidium bromide staining of products separated electrophoretically on agarose/TBE gels (Maniatis et al., 1982). (Additionally, in the case of the Hind III digest, a Southern blot of the products was also probed with nick-translated 5S DNA and a single band was observed—as expected for a complete digest, because there is a Hind III site within the spacer sequence of the 5S RNA genes of Xenopus laevis.) Electrophoretically separated digestion products were blotted to nitrocellulose and treated under stringent conditions (Southern, 1975; Maniatis et al., 1982). Prehybridizations were in 4X SSC, 4X Denhardt's, 10 mM EDTA, 0.5% SDS, 100 μg/ml sonicated salmon sperm DNA, and 5% dextran sulfate, at 65°C–68°C for 4–6 hr. Hybridizations were in 4X SSC, 4X Denhardt's, 10 mM EDTA, 0.5% SDS, 100 μg/ml sonicated salmon sperm DNA, 10% dextran sulfate, and 10^7–10^8 cpm of nick-translated puc3a1.b insert DNA (10^6 cpm/μg) at 65°C–68°C for 15 hr. The filters were washed twice for 10 min at room temperature in 2X SSC/0.1% SDS, then three times for 60 min at 65°C in 0.1X SSC/0.1% SDS). Identically digested puc3a1.b insert DNA electrophoresed on the same gels as genomic DNA and transferred to the same nitrocellulose filters served as a standard for quantitation of these blots. Quantitation was by densitometric scanning of autoradiograms with a DU8B spectrophotometer (Beckman).

Northern Blot Hybridization Analysis of RNA from Staged Oocytes and Embryos

Oocytes were staged according to Dumont (1972) as follows: Stage I and II: translucent–white, ≤400 μ. Stage III: light brown–blackish-brown, 450–600 μ. Stage IV: animal and vegetal hemispheres differentiated, 600–1000 μ. Stage V: hemispheres clearly delineated, 1000–1200 μ. Stage VI: white equatorial band present, 1200–1300 μ. Embryos were staged according to Nieuwkoop and Faber (1956): Fertilized egg = Stage 1; blastula = Stages 8/9; neurula = Stage 19; swimming tadpole = Stages 40/41.

Total cellular RNA was isolated from staged oocytes and embryos, and from adult liver and a Xenopus adult kidney cell line (cell line was a gift from Dr. D. Shapiro, University of Illinois; kidney cell RNA was prepared from actively growing cells by Dr. M. Perry, this laboratory) according to the guanidinium isothiocyanate/CsCl procedure (Maniatis et al., 1982). ^{3}H-labeled nuclear RNA purified from HeLa cells (a gift from H. Sive, this laboratory) was added to each pool of oocytes/embryos just prior to homogenization. Recoveries of RNA were estimated relative to recovery of ^{3}H-RNA from each preparation. ^{3}H-RNA blotted to nitrocellulose (same amount as used for marker) did not produce a signal on autoradiograms of Northern blots. RNA samples were separated electrophoretically and transferred to nitrocellulose as previously described (Maniatis et al., 1982). Filters were probed under the conditions used for Southern blot hybridizations, above, except that prehybridizations did not contain any dextran sulfate. Nick-translated puc3a1.b insert DNA served as a probe. Hybridization signals were quantitated as for Southern blot analyses, above.

Sequence Homology Searches

puc3a1.b cDNA sequence was compared to that of various 5S RNA (Miller et al., 1978) and tRNA (Galli et al., 1981) genes and to nucleotide sequences of the lambda *cro* (Roberts et al., 1977), and *cI* genes (Sauer, 1978) with the help of the NUCALN homology search program devised by D. J. Lipman and W. J. Wilbur (NIADDK, NIH).

Acknowledgments

We thank Dr. Douglas Melton and his colleagues for their generous gift of the cDNA library as well as for stimulating discussions and kind hospitality. We are grateful to Dr. Michael Hunkapiller for peptide sequencing and Wilder Fulford, Steve Loukin, and Paul Martin for help in oligonucleotide syntheses and purification. We also thank Drs. Richard Young, Mark Dworkin, Eva Dworkin, Nathaniel Heintz, and members of the Roeder laboratory for helpful discussions. This research was supported by National Institutes of Health grant R01 CA34891. A. M. G. is a recipient of a fellowship from the NIH Medical Scientist Training Program under grant GM07200 to Washington University Medical School, St. Louis, Missouri.

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

Received September 19, 1984

References

Agarwal, K. L., Brunstedt, J., and Noyes, B. (1981). A general method for detection and characterization of an mRNA using an oligonucleotide probe. J. Biol. Chem. *256*, 1023–1028.

Bieker, J. J., and Roeder, R. G. (1984). Physical properties and DNA-binding stoichiometry of a 5S gene-specific transcription factor. J. Biol. Chem. *259*, 6158–6164.

Bisbee, C. A., Baker, M. A., Wilson, A. C., Hadji-Azini, I., and Fischberg, M. (1977). Albumin phylogeny for clawed frogs (Xenopus). Science *195*, 785–787.

Bogenhagen, D. F., Sakonju, S., and Brown, D. D. (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.

Davis, R. W., Botstein, D., and Roth, J. R. (1980). A Manual for Genetic Engineering: Advanced Bacterial Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Dignam, J. D., Martin, P. L., Shastry, B. S., and Roeder, R. G. (1983). eukaryotic gene transcription with purified components. Meth. Enzymol. *101,* 582–598.

Dumont, J. (1972). Oogenesis in *Xenopus laevis* (Daudin); stages of oocyte development in laboratory maintained animals. J. Morph. *136,* 153–180.

Engelke, D. R., Ng, S.-Y., Shastry, B. S., and Roeder, R. G. (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell *19,* 717–728.

Ford, P. J. (1971). Non-coordinated accumulation and synthesis of 5S ribonucleic acid by ovaries of *Xenopus laevis*. Nature *233,* 561–564.

Galli, G., Hofstetter, H., and Birnstiel, M. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature *294,* 626–631.

Golden, L., Schafer, U., and Rosbash, M. (1980). Accumulation of individual pA^+ RNAs during oogenesis of Xenopus laevis. Cell *22,* 835–844.

Hofstetter, H., Kressman, A., and Birnstiel, M. L. (1981). A split promoter for a eucaryotic tRNA gene. Cell *24,* 573–585.

Honda, B. M., and Roeder, R. G. (1980). Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. Cell *22,* 119–126.

Hunkapiller, M., Henwick, R., Dreyer, W., and Hood, L. (1983a). High sensitivity sequencing with a gas-phase sequenator. Meth. Enzymol. *91,* 399–413.

Hunkapiller, M., Lujan, E., Ostrander, F., and Hood, L. (1983b). Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Meth. Enzymol. *91,* 227–36.

Korn, L. (1982). Transcription of *Xenopus* 5S ribosomal RNA genes. Nature *295,* 101–105.

Lassar, A. B., Martin, P. L., and Roeder, R. G. (1983). Transcription of class III genes: formation of preinitiation complexes. Science *222,* 740–748.

Maniatis, T., Fritsch, E., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65,* 499–560.

Miller, J. I., Cartwright, E. M., Brownlee, G. G., Federoff, N. V., and Brown, D. D. (1978). The nucleotide sequence of oocyte 5S DNA in Xenopus laevis. II. The GC-rich region. Cell *13,* 717–725.

Myoshi, K., Miyaka, T., Hozumi, T., and Itakura, K. (1980). Solid phase synthesis of polynucleotides. Nucl. Acids Res. *8,* 5473–5505.

Nevins, J. R. (1983). The pathway of eukaryotic mRNA formation. Ann. Rev. Biochem. *52,* 241–245.

Newport, J., and Kirschner, M. (1982). A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. Cell *30,* 687–696.

Nieuwkoop, P. D., and Faber, J. (1956) Normal Table of *Xenopus laevis* (Daudin) (Amsterdam: North Holland Publishing Company).

Pelham, H., and Brown, D. D. (1980). A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc. Nat. Acad. Sci. USA *77,* 4170–4174.

Pelham, H., Wormington, W. M., and Brown, D. D. (1981). Related 5S RNA transcription factors in *Xenopus* oocytes and somatic cells. Proc. Nat. Acad. Sci. USA *78,* 1760–1764.

Picard, B., and Wegnez, M. (1979). Isolation of a 7S particle from *Xenopus laevis* oocytes: a 5S RNA–protein complex. Proc. Nat. Acad. Sci. USA *76,* 241–245.

Pieler, T., and Erdmann, V. (1983). Isolation and characterization of a 7S RNP particle from mature *Xenopus laevis* oocytes. FEBS Lett. *157,* 283–287.

Roberts, T., Shimatake, H., Brady, C., and Rosenberg, M. (1977). Sequence of *cro* gene of bacteriophage lambda. Nature *270,* 274–275.

Roeder, R. G., Lee, D. C., Honda, B. M., and Shastry, B. S. (1981). Analysis of eukaryotic gene transcription *in vitro*. ICN–UCLA Aymposium on Developmental Biology Using Purified Genes, D. Brown and D. Fox, eds. (New York: Academic Press), pp. 429–446.

Rosbash, M., and Ford, P. (1974). Polyadenylic acid-containing RNA in *Xenopus laevis* oocytes. J. Mol. Biol. *85,* 87–101.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. USA *74,* 5463–5467.

Sauer, R. (1978). DNA sequence of the bacteriophage λ cI gene. Nature *276,* 301–302.

Segall, J., Matsui, T., and Roeder, R. G. (1980). Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III. J. Biol. Chem. *255,* 11986–11991.

Sgaramella, V., and Khorana, H. G. (1972). CXII. Total synthesis of the structural gene for an alanine transfer RNA from yeast. Enzymic joining of the chemically synthesized polydeoxynucleotides to form the DNA duplex representing nucleotide sequence 1 to 20. J. Mol. Biol. *256,* 7990–7997.

Shastry, B. S., Ng, S.-Y., and Roeder, R. G. (1982). Multiple factors involved in the transcription of class III genes in *Xenopus laevis*. J. Biol. Chem. *257,* 12979–12986.

Shastry, B. S., Honda, B. M., and Roeder, R. G. (1984). Altered levels of 5S gene-specific transcription factor (TFIIIA) during oogenesis and embryonic development of *Xenopus laevis*. J. Biol. Chem., in press.

Smith, D. R., Jackson, I. J., and Brown, D. D. (1984). Domains of the positive transcription factor specific for the Xenopus 5S RNA gene. Cell *37,* 645–652.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98,* 503–518.

Suggs, S., Hirose, T., Miyake, T., Kawashima, E., Johnson, M., Itakura, K., and Wallace, B. (1981). Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences. ICN–UCLA Symposium on Developmental Biology Using Purified Genes, D. Brown and D. Fox, eds. (New York: Academic Press), pp. 683–693.

Wormington, W. M., and Brown, D. D. (1983). Onset of 5S RNA gene regulation during *Xenopus* embryogenesis. Dev. Biol. *99,* 248–257.

Wormington, W. M., Schlissel, M., and Brown, D. D. (1983). Developmental regulation of the *Xenopus* 5S RNA genes. Cold Spring Harbor Symp. Quant. Biol. *47,* 879–884.