Nucleic Acids Research, Vol. 18, No. 1

Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P

Madeline Baer, Timothy W.Nilsen+, Christine Costigan and Sidney Altman*
Department of Biology, Yale University, New Haven, CT 06520, USA

Received October 2, 1989; Revised and Accepted November 15, 1989

ABSTRACT

The gene coding for H1 RNA, the RNA component of human RNase P, has been isolated and characterized from a human genomic DNA library. The sequence corresponding to the mature H1 RNA is almost identical to that previously identified using H1 RNA and a cDNA clone corresponding to it. The nucleotide sequence of the genomic clone contains an array of potential transcriptional control elements, some characteristic of transcription by RNA polymerase III and some characteristic of RNA polymerase II, as is also the case for U6 and certain other small stable RNAs. The transcription *in vitro* of the genomic clone shows that the gene is functional and is transcribed by RNA polymerase III. Southern hybridization analysis indicates that there is very likely only one copy of the gene for H1 RNA in the human genome.

INTRODUCTION

RNase P is an endoribonuclease that cleaves tRNA precursor molecules to form the mature 5′ termini of their tRNA sequences. In both prokaryotic and eukaryotic organisms the enzyme exists as a ribonucleoprotein (1,2) and the RNA subunit from several of the prokaryotic enzymes has been shown to be catalyticaly active alone *in vitro*. Thus far, however, the eukaryotic enzymes are only active *in vitro* in the form of a ribonucleoprotein complex (3). The sequence of the RNA component of human RNase P from Hela cells, H1 RNA, has been determined by direct sequencing of the H1 RNA and determination of the sequence of a cDNA clone corresponding to this RNA (4). To verify this sequence, and also to obtain information concerning the control of transcription of the gene for H1 RNA, we sought a clone containing a genomic copy of the H1 RNA gene.

We have now isolated and sequenced a gene coding for the RNA portion of human RNase P and its flanking regions from the DNA of human spleen. The sequence of the transcribed portion of the gene is nearly identical to the cDNA sequence of the H1 RNA molecule from Hela cells previously published (4). The gene is transcribed *in vitro* by RNA polymerase III. However, it appears to have transcriptional control elements characteristic of those used by both RNA polymerases II and III. It, therefore, may be similar in its transcriptional control to the genes of several other eukaryotic small stable RNAs previously characterized including U6, 7SK, 7SL, MRP, and the Epstein-Barr Virus Small RNAs. (5−9)

MATERIALS AND METHODS

Analysis of Human DNA

Aliquots of DNA (20 μg) from human spleen (a gift of Dr. B. Forget, Yale University) were cut with a variety of restriction enzymes and subjected to electrophoresis on 0.8% agarose gels. After transfer (28) to either nitrocellulose or nylon (Zeta-Probe) filters, hybridizations were carried out using ^{32}P-labeled cDNA that codes for H1 RNA of RNase P from HeLa cells (4). The cDNA was labeled with ^{32}P using a random priming kit purchased from Boehringer-Mannheim. Hybridizations to DNA bound to the Zeta-Probe membrane were performed with aqueous solutions containing 10% dextran sulphate, 1.5× SSPE (0.27 M NaCl, 15 mM sodium phosphate, pH 7.0, 1.5 mM EDTA), 1 % SDS, and 0.5% non-fat dry milk at 68° C. The final wash of the Zeta -probe membrane was done at 50° C in the presence of 0.1× SSC (15 mM NaCl, 1.5 mM sodium citrate) and 1% SDS. Hybridizations to DNA bound to the nitrocellulose membrane were done at 42° C using a solution containing 50% formamide, 10% dextran sulphate, 5× SSC, 2.5× Denhardt's buffer (29), and 0.2% SDS. The final wash was done at 65° C in 0.2× SSC. In all cases sonicated, denatured salmon sperm DNA was included in the hybridization mixes.

Isolation of the Gene for H1 RNA

DNA (1 μg) from human spleen was digested with BamH I and ligated to 500 ng of λEMBL3 arms (Promega). The ligation reaction was packaged using Gigapack Gold packaging mix (Stratagene), and the phage were grown using *E.coli* strain KW251 (Promega) as a host. Approximately 10^6 pfu were obtained per mg of human DNA.

The phage were plated without amplification. After transfer to nitrocellulose filters and denaturation of the bound DNA (29), the plaques were screened for the presence of the gene for H1 RNA using the ^{32}P-labeled cDNA for Hela cell H1 RNA as a probe.

Phage stocks carrying the gene for H1 RNA were prepared and their DNAs were extracted according to the procedure of Kaslow (30). After digestion with a variety of restriction enzymes

* To whom correspondence should be addressed

+ Present address: Department of Molecular Biology and Microbiology, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

and subsequent analysis of the phage DNA by agarose gel electrophoresis, fragments containing the H1 RNA gene were identified following Southern transfer and hybridization to the ^{32}P-labeled probe DNA.

Subcloning and DNA Sequencing

The phage DNA was cut with Sal I (a unique Sal I site is located next to the BamH I site in the λEMBL3 arms), and a fragment of DNA of approximately 15 kb was produced and cloned into pUC 19 (31) at the Sal I site of the polylinker. This plasmid was named pMBH1. The insert in pMBH1 was then further subcloned. This DNA was digested with Sma I and EcoR I and the resulting fragment of 831 bp which contained the gene for H1 RNA was cloned into pUC 19 at the Sma I and EcoR I sites of the polylinker. The new plasmid was named pMBH2. It contained 835 bp of the genomic DNA sequence. The sequence of the inserted DNA was determined using the Sequenase™ method and kit (United States Biochemical). The oligonucleotide primers used were the minus 40 Universal Primer and the Reverse Primer (New England Biolabs), which hybridize to the vector on either side of the inserted DNA, and a primer with the sequence 5′-GCATCTCCTGCCCAGTCTGACCTCG-3′, which hybridizes within the coding region of the H1 RNA gene to nucleotides 163−187. The fragment of 835 bp was further subcloned by digestion with Kpn I and either EcoR I or Pst I. Each of these fragments, which contained part of the gene for H1 RNA, was cloned into pUC 19 at the appropriate sites in the polylinker. The sequence of the inserted DNAs was then determined as decribed above using both the Universal and Reverse primers. The sequence of the DNA located beyond the Sma I and EcoR I sites was determined using the DNA from the pUC 19 clone containing the 15 kb fragment of genomic DNA carrying the H1 RNA gene and three primers that hybridize to nucleotides either in or beyond the coding region of the H1 RNA gene. These three primers are 5′-GGGCCCGCGATTCCT-TGGA-3′, which is complementary to nucleotides −182 to −200, 5′-TCGTGGCCCCACTGATGAGCTT-3′, which is complementrary to nucleotides 14−35 of the coding region, and 5′-TGGAACAGA CTCACGGCCAG-3′, which hybridizes to nucleotides 242−261 of the antisense strand of the DNA in the coding region.

Subclone of the DNA fragment of 835 bp

The plasmid containing the Sma I-EcoR I fragment of genomic DNA (835 bp) was digested with Ssp I, Hind III and EcoR I. The Ssp I-Hind III fragment containing the H1 RNA gene that arises from the triple digest was cloned into pUC 19 at the unique Ssp I and Hind III sites and the new plasmid was called pMBH3. This plasmid contained 706 bp of the genomic DNA sequence.

Preparation of cell extracts

S100 extracts of Hela cells were prepared as previously described (32). To prepare whole cell extracts, Hela cells were concentrated by centrifugation and washed twice with ice-cold phosphate buffered saline (Gibco). Following one additional wash in 10mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, 10 mM Hepes, pH 7.9 (buffer A; 33), the cells were resuspended in one packed cell volume of buffer A and allowed to swell for 10 min. Following swelling, the cells were disrupted by Dounce homogenization until most nuclei were broken, as monitored by phase contrast microscopy. The homogenate was then made 200 mM in KCl by addition of 1M KCl and further homogenized by 5 strokes of the Dounce homogenizer before centrifugation at 26,000 g for 10 min. The resultant supernatant was dialysed for 4 hours against 100 volumes of 100mM KCl, 0.5 mM DTT, 0.2mM EDTA, 20 % glycerol, 20 mM Tris·HCl, pH 7.9. Following dialysis the homogenate was centrifuged at 10,000 g for 15 min and the resultant supernatant was divided into aliquots and stored at −70° C.

Transcription reactions *in vitro*

Transcription reactions *in vitro* were carried out in a final volume of 25 μl. The mixtures contained 15 μl of the indicated extract, 5 mM $MgCl_2$, 80 mM KCl, 0.5 mM DTT, 20 mM creatine phosphate, 0.5 mM each of ATP, UTP, CTP, 50 μM GTP, 10 μCi (α-^{32}P GTP), 12mM Tris·HCl, pH 7.9 and DNA templates as indicated in the figure legends. Following a one hour incubation at 30° C reaction mixtures were diluted to 250 μl in 250 mM NaOac, 1 mM EDTA, 0.25 % SDS, 20 mM Tris·HCl, pH 7.5, digested with 100 μg/ml proteinase K for 30 min at 37° C and extracted with phenol/chloroform (29). Reaction products were analyzed on 8M urea, 5 % polyacrylamide gels. Following electrophoresis, labeled transcripts were visualized by autoradiography.

Isolation of RNA transcribed *in vivo*

Whole cell extracts of HeLa cells, made as described above, were deproteinized by proteinase K treatment and extraction with phenol and chloroform. The total cellular RNA obtained was precipitated with ethanol prior to its hybridization with oligonucleotides and cleavage by RNase H.

RNase H cleavage reactions

RNA isolated either from transcription reactions *in vitro* or Hela cell extracts was individually hybridized to two oligonucleotides, the sequences of which are given above, and which are complementary to nucleotides 14−35 or 163−187 of H1 RNA. Cleavage by RNase H was performed as described by Vournakis et al. (34). Reaction products were fractionated on 5% denaturing acrylamide gels and transferred to Gene Screen (NEN).

Blot hybridization

Following isolation from whole cell extracts, hybridization to oligonucleotides and cleavage with RNase H, RNA was fractionated on 5% denaturing acrylamide gels and transferred to Gene Screen (NEN). Probes were prepared by labeling oligonucleotides complementary to nucleotides 14−35 and 163−187 of H1 RNA with γ-^{32}P-ATP and polynucleotide kinase. The hybridization mix contained 2×10^7 cpm of mixed oligonucleotide probe in 10 mg/ml BSA, 1mM EDTA , 7% SDS, 0.5M sodium phosphate, pH 7.2. Hybridizations were carried out at 37° C for 16 hours. Washes were done at 37° C in 5× SSC containing 0.1% SDS.

Mapping of the chromosomal position of the gene for H1 RNA

Initial mapping experiments were performed using DNA mapping panels as described by Bentley et al. (21) and references therein. The cDNA clone isolated by Barkiewickz et al. (4) was labeled with ^{32}P and used as the probe in these experiments. More definitive mapping of the chromosomal position was done using the technique of hybridization *in situ* developed by D. Ward and P. Lichter (personal communication). The fragment of spleen DNA, 15 kb in length, which had been subcloned into pUC 19, was used as the probe.

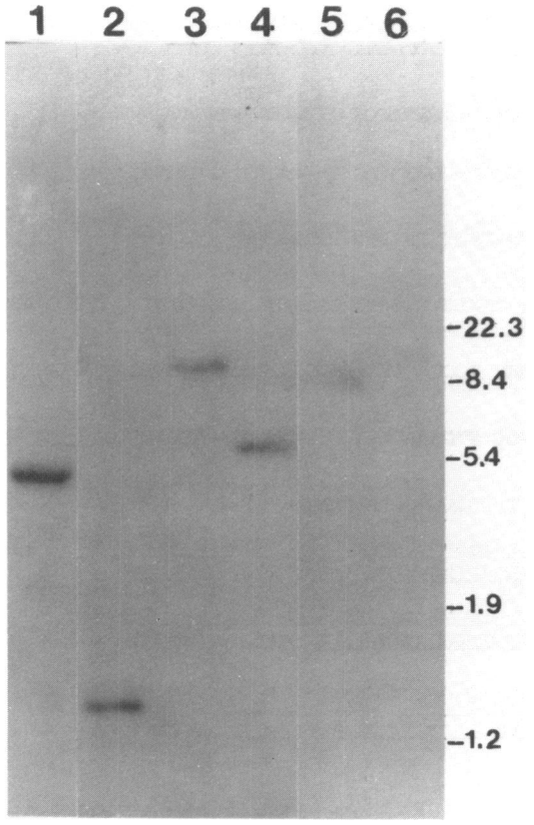

Figure 1. Southern hybridization of H1 cDNA to human spleen DNA cut with various restriction enzymes. Human spleen DNA was cut with a variety of restriction enzymes, subjected to electrophoresis on an 0.8% agarose gel, transferred to nitrocellulose and hybridized under stringent conditions to a ^{32}P-labeled cDNA fragment corresponding to H1 RNA of Hela cell RNase P (see Materials and Methods.). The enzymes used were: lane 1, EcoR I and Xba I; lane 2; Ava I and Xba I; lane 3, Xho I and EcoR I; lane 4, BamH I and EcoR I; lane 5, EcoR I; lane 6, BamH I. Size markers are as shown.

RESULTS

Cloning of the gene for H1 RNA from human DNA

DNA from human spleen was cut with a variety of restriction enzymes, fractionated by agarose gel electrophoresis, and analyzed by Southern transfer and hybridization to the cDNA corresponding to the gene for H1 RNA of Hela cell RNase P (4). A single BamH I fragment of about 15 kb that hybridized to the probe was identified as shown in Fig. 1, lane 6.

To isolate the gene coding for H1 RNA, the human spleen DNA was cut with BamH I and ligated into λEMBL3 arms to produce a human spleen genomic library of BamH I fragments between 9 and 23 kb in size. Approximately 10^6 plaques were screened using the cDNA corresponding to H1 RNA as a probe (see Materials and Methods). Twelve positive plaques were obtained. Large scale lysates of four of these phage were prepared and their DNA was extracted and cut with various restriction enzymes. All four DNAs had the expected BamH I fragment of 15 kb. All four DNAs also carried the gene for H1 RNA on a single Pst I fragment of about 2 kb and a single Sma I-Sac I fragment of approximately 900 bp. One of these phage DNAs (the DNA of isolate 9), was chosen for further study. The DNA fragment of 15 kb from this phage was subcloned into pUC 19 to make pMBH1. A Sma I-EcoRI fragment of approximately 830 bp which contained the gene was subcloned separately into pUC 19 to make pMBH2 (see Materials and Methods).

Sequence of the gene for H1 RNA and its surrounding regions

The sequence of the gene for H1 RNA was determined using pMBH1 and pMBH2, the recombinant plasmids described above (see Materials and Methods). In comparison with the previously published sequence of H1 RNA, the coding region of the gene differs from the cDNA sequence at three positions (Fig. 2). It has one extra A residue after position 2 of the cDNA sequence. Position 129 of the genomic sequence is changed from A in the cDNA to G in the genomic sequence and position 324 is changed from C in the cDNA to T in the genomic sequence. We attribute these differences to polymorphisms between the Hela cell DNA and human spleen DNA. The two transitions found in the H1 RNA coding sequence of the spleen DNA do not affect the proposed secondary structure of the H1 RNA (4). The additional A residue close to the 5′ end of the coding region still allows the 5′ and 3′ ends of the H1 RNA to form a based paired region. However, this extra A residue must bulge out of the proposed helical region if the stem structure is to be preserved.

Sequences characteristic of genes transcribed by RNA polymerase III have been identified in the gene for H1 RNA. Within the coding sequence of the gene, at position 269−278, is the sequence 5′-AGTTCAATGG-3′ which matches the box A consensus sequence (5′-RRYNNARYGG-3′) found as part of RNA polymerase III promoters (10). However, a box B consensus sequence was not present. At the 3′ end of the gene there is a run of 5 T residues which could serve as a signal for termination of transcription by RNA polymerase III (11,12).

In addition to the sequences which match the consensus sequences for control elements for transcription by RNA polymerase III, the gene for H1 RNA contains in its 5′ flanking region several sequences which match the consensus sequences of control elements for RNA polymerase II. These include the sequence 5′-TATAA-3′ located at position −30, which matches the TATA box control element sequence (13); the sequence 5′-TCACCATAAAC-3′ located at position −68, which matches the consensus sequence 5′-TYACCNTAAC-3′ for the proximal sequence element (14,15); the sequence 5′-ATTTGCAT-3′ at position −97, which is identical to the distal sequence element (octamer sequence; 13); and 5 short, GC rich sequences, located upstream of the distal sequence element, which could be binding sites for the Sp1 transcription factor (16). One of these Sp 1 elements is separated from the octamer sequence by only 6 bases. The orientation of the octamer sequence is identical to that found in the genes for U6 RNA, MRP RNA, U4 RNA, and 7SK RNA and opposite to that seen in the genes for the U1 and U2 RNAs (5,8,17−19).

At the 3′ end of the H1 RNA gene, following the 3′ terminal stretch of T residues are two G residues followed directly by a long run of A residues and an Alu sequence (20). The Alu sequence is directly followed by a poly A sequence at its 3′ end.

Copy number and chromosomal location of the H1 RNA gene

Southern hybridization analyses of single and double digests of the human spleen DNA show a single band of hybridization indicating that the gene for H1 RNA may be present in a single copy, (Figure 1). On occasion, with some enzymes we have observed hybridization to two bands in restriction digests. Even though these results may be due to partial digests, nonspecific hybridization or to hybridization to a truncated gene, we cannot

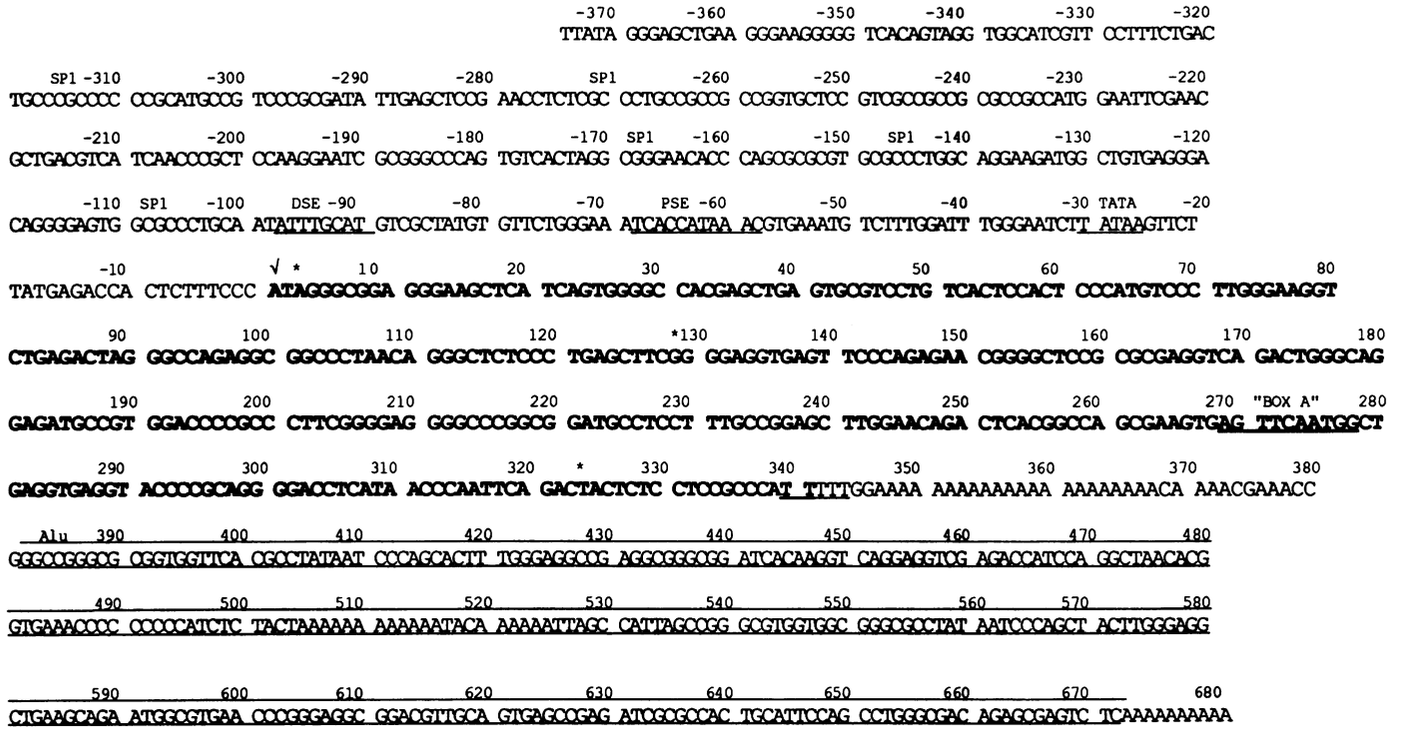

Figure 2. Sequence of the gene for H1 RNA. The region corresponding to the cDNA sequence for H1 RNA is shown in bold type. The first base of H1 RNA is marked with a check. Bases which differ from the Hela cell H1 RNA sequence are marked by asterisks. Sequences which match consensus sequences for transcriptional control elements are as marked and include a box A sequence (Box A), a TATA box (TATA), a proximal sequence element (PSE), a distal sequence element (DSE), and several Sp1 elements (Sp1). The terminal T residues at the 3′ end of the gene are underlined. The Alu sequence is under and over lined.

rigorously exclude the possibility that the gene for H1 RNA is actually present in the human genome in two copies. Digestion of the DNAs of the recombinant phage with Sma I and Sac I and analysis of the digests by Southern hybridization did show that in all four DNAs the gene for H1 RNA was carried on a unique fragment of approximately 900 bp.

Preliminary chromosome mapping using panels of DNA from cells carrying mouse-human hybrid chromosomes (21) localized the gene to either one of two chromosomes, 11 or 14. Subsequently, using the high resolution technique of Ward and Lichter (personal communication), the gene has been mapped to chromosome 14. It resides on the q arm just below the centromere.

Transcription *in vitro* of the gene for H1 RNA

Various transcription experiments were carried out *in vitro* to determine whether or not the gene for H1 RNA was functional and to determine the type of RNA polymerase that might be involved in its transcription. Both pMBH1 and pMBH2 were transcribed *in vitro* using S100 and whole Hela cell extracts. Plasmid DNA containing the gene for 5S rRNA was also transcribed in the same extracts as an internal control. Synthesis of a transcript about 340 nucleotides long, approximately the size of H1 RNA present in RNase P (4), was observed when templates containing the gene for H1 RNA were incubated with either cell extract. Data for pMBH2 are shown in Figs. 3, 4 and 5. The level of synthesis of this transcript was the same when plasmid DNA carrying the genomic DNA fragment of 15 kb (pMBH1) or plasmid DNA containing the piece of genomic DNA of 835 bp (pMBH2) with only 229 bases of 5′ flanking sequence was used as a template (data not shown).

To establish that the transcript of about 340 nucleotides was indeed H1 RNA, the RNA produced *in vitro* was digested with RNase H in the presence of one or the other of two oligonucleotides complementary to H1 RNA (Figure 3A). These two oligonucleotides hybridize to nucleotides 14−35 and 163−187 of the RNA (Materials and Methods). Similar analyses were also performed on total RNA made *in vivo* which was isolated from Hela cell extracts (Figure 3B). In these latter experiments H1 RNA was identified by blot hybridization using oligonucleotide probes complementary to the H1 RNA sequence. Both RNAs yielded a nearly identical spectrum of digestion products. Cleavage by RNase H of transcripts hybridized to the oligonucleotide complementary to nucleotides 14−35 of the H1 RNA gene produced, as expected, a large fragment of approximately 318 bases (the smaller fragment ran off the gel), while cleavage of the transcripts hybridized to the oligonucleotide complementary to nucleotides 163−187 of the H1 RNA gene produced a broad band. This band corresponded in size and intensity to the two fragments of around 170 nucleotides which should be produced in this particular RNase H cleavage reaction (Figure 3, lanes 2 and 3 in panels A and B).

Although the analysis with RNase H does not provide definitive information regarding the actual initiation and termination sites of transcription, the results do show that the transcript synthesized *in vitro* using the cloned gene as a template is H1 RNA, that transcription is accurate, and that the transcript made is identical

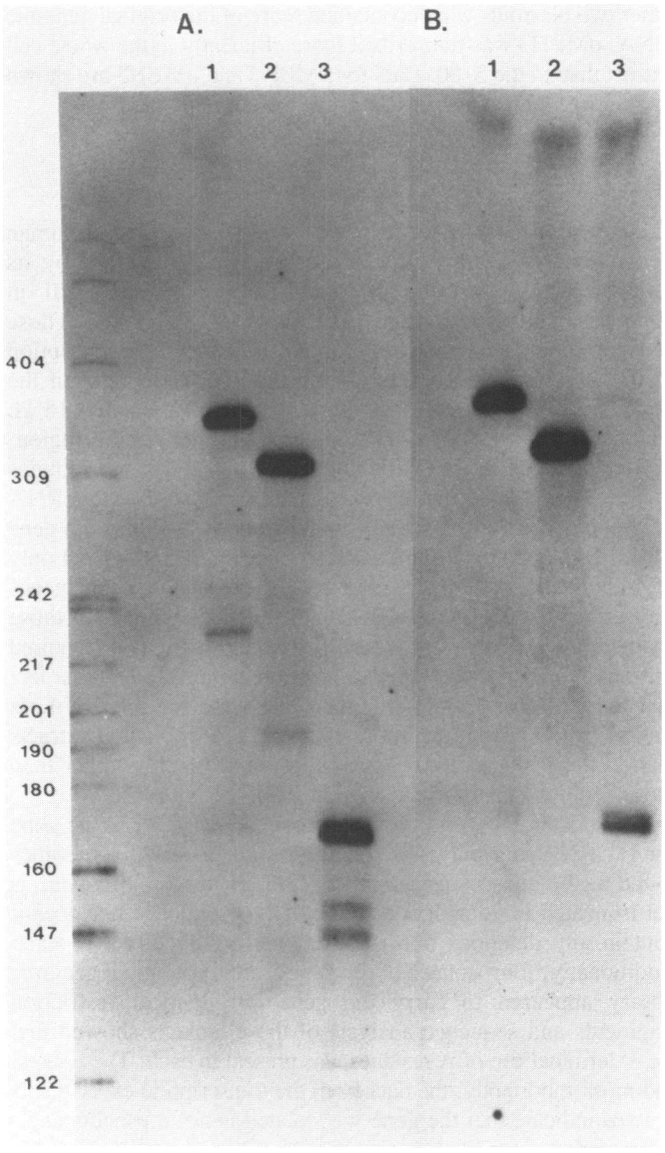

Figure 3. Comparison of transcription products made *in vitro* and *in vivo* from the gene for H1 RNA. Panel A: Autoradiogram of an RNA blot of the experiment in which 500 ng of supercoiled plasmid DNA of pMBH2 (see Materials and Methods) were incubated in a whole cell extract of HeLa cells as described in Materials and Methods. Following deproteinization, the reaction products were divided into three aliquots, hybridized to various oligonucleotides and digested with RNase H. The reaction products were then separated on denaturing polyacrylamide gels and transferred to Gene Screen (NEN) as described in Materials and Methods. Panel B: Autoradiogram of an RNA blot in which aliquots of whole cell RNA from HeLa cells (1,2 and 3 μg respectively in lanes 1, 2 and 3) were treated as described above for the transcripts made *in vitro*. The part of the filter corresponding to panel A was exposed directly to film, while the portion of the filter corresponding to panel B was hybridized with labeled oligonucleotide as described in Materials and Methods and then exposed to film. For the experiments shown in both panels, the digestions with RNase H were carried out (lanes 1) in the absence of oligonucleotide; (lanes 2) in the presence of 200 ng of an oligonucleotide complementary to nucleotides 14–35 of H1 RNA, and (lanes 3) in the presence of 200 ng of an oligonucleotide complementary to nucleotides 163–187 of H1 RNA. The sizes of single stranded DNA markers are indicated in the figure.

or nearly so in size to the transcript made *in vivo*. We also conclude that the mature form of H1 RNA does not appear to be derived from a larger precursor molecule.

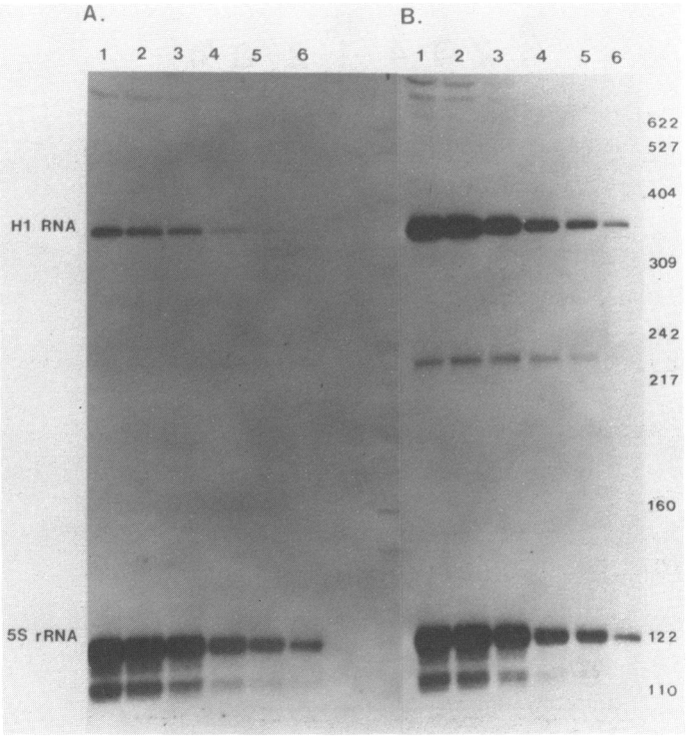

Figure 4. Inhibition of transcription of the gene for H1 RNA by α-amanitin. 250 ng each of plasmid DNA (pMBH2) (see text and Materials and Methods) or plasmid DNA containing a gene for *Xenopus borealis* 5S rRNA (pXBS201, 35) were incubated with an S100 extract (panel A) or a whole cell extract of HeLa cells (panel B), in the presence of increasing concentrations of α-amanitin. For both panels: lane 1, no addition; lane 2, 2 μg /ml; lane 3, 10 μg /ml; lane 4, 50 μg /ml; lane 5,100 μg/ml; lane 6, 200 μg/ml. Reactions were assembled and analyzed as described in Materials and Methods. The transcript of approximately 225 nucleotides seen in the whole cell extracts has not been rigorously identified, but based upon its sensitivity to digestion with RNase H in the presence of oligonucleotides complementary to H1 RNA, it is likely to be a prematurely terminated version of H1 RNA. The sizes of single-stranded DNA markers are indicated in the figure.

RNA polymerase II or RNA polymerase III?

To ascertain the nature of the polymerase which transcribes the gene for H1 RNA, transcription reactions were carried out *in vitro* in the presence of increasing concentrations of α-amanitin. As an internal control, these reactions contained an authentic RNA polymerase III template corresponding to a gene for 5S rRNA from *Xenopus borealis*. H1 RNA and 5S rRNA synthesis was assayed in both S100 and whole cell extracts (Figure 4). Synthesis of both transcripts showed an identical sensitivity to inhibition by α-amanitin. That is, the transcription of both RNA molecules was not completely inhibited even at 200 μg/ml concentrations of α-amanitin. These results indicate that the H1 RNA gene is transcribed by RNA polymerase III. Interestingly, H1 RNA, unlike 5S rRNA, was synthesized much more efficiently in the whole cell extract than in the S100 extract, a result which suggests that the transcription of H1 RNA may require factors in addition to those utilized for the transcription of 5S rRNA.

To delineate further the extent of the 5′ flanking region of the H1 gene required for efficient transcription, the gene was cut at an Ssp I site at nucleotide −98, adjacent to the 5′ end of the octamer sequence. The resulting fragment, after recloning to make plasmid pMBH3, contains 706 bp of the initial genomic

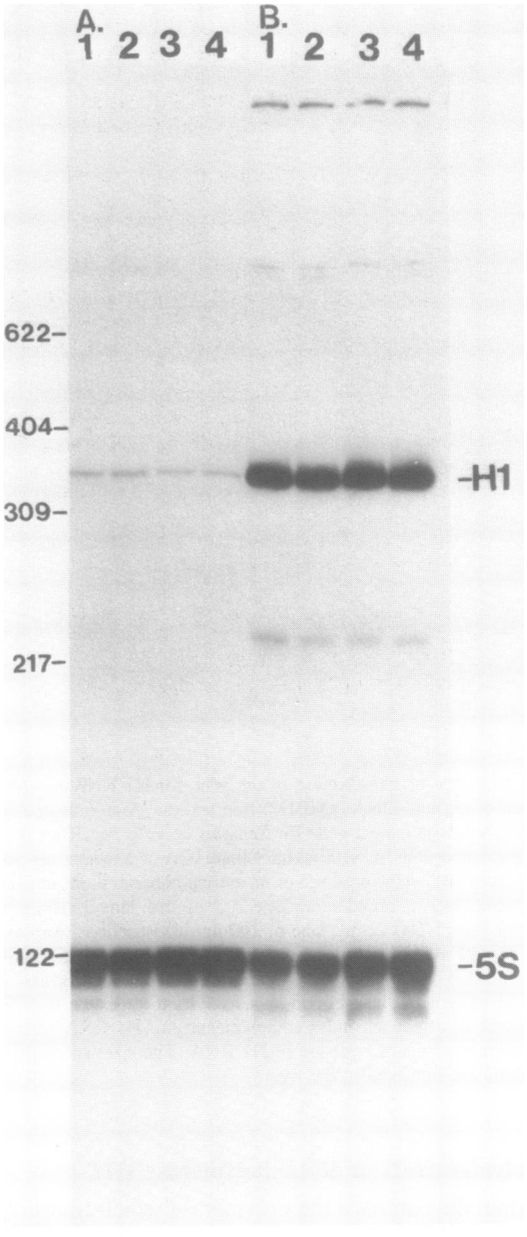

Figure 5. Transcription *in vitro* of an H1 RNA gene lacking potential Sp1 binding sites. Plasmid DNA containing inserts of human DNA carrying the H1 RNA coding sequence, with either 229 nucleotides of 5′ flanking genomic DNA, (pMBH2; lanes 1 and 2 in both panels) or 100 nucleotides of 5′ flanking genomic DNA, (pMBH3; lanes 3 and 4 in both panels) were incubated with an S100 extract of HeLa cells (Panel A) or whole cell extract of Hela cells (panel B). Lanes 1 and 3 contained 250 ng of H1 RNA plasmid DNA while Lanes 2 and 4 contained 350 ng of the template. All reactions contained 250 ng of plasmid DNA carrying the *Xenopus borealis* 5S rRNA gene described in the legend to Figure 4. Reaction mixtures were assembled and analyzed as described in Materials and Methods. The sizes of single stranded DNA markers are indicated.

fragment with only 100 nucleotides of 5′ flanking DNA included. This fragment still contains the intact octamer sequence. However, all of the putative Sp1 sequences have been deleted.

pMBH3 DNA was transcribed *in vitro* in both the S100 and whole cell extracts as described above. The gene for H1 RNA was transcribed just as efficiently from this DNA, in which only 100 nucleotides of 5′ flanking genomic DNA are present, as it was from the DNA of plasmid pMBH2 that contains more of the 5′ flanking region (Figure 5). As was also the case with the other two plasmids which contained more of the original genomic DNA, pMBH3 was transcribed more efficiently in the whole cell extract than in the S100. Data for pMBH2 and pMBH3 are shown in Figure 5. Data for pMBH1 are not shown.

DISCUSSION

The gene for H1 RNA which we have isolated from a human genomic DNA library is a functional gene as judged by its efficient transcription *in vitro* by RNA polymerase III in comparison to transcription of a gene for 5S rRNA. Those sequence elements which are required for efficient transcription of the gene for H1 RNA lie within the 706 nucleotides of the genomic fragment subcloned from the initially isolated 15 kb fragment. This includes 341 nucleotides of the coding region, 100 nucleotides in the 5′ flanking region and 265 nucleotides in the 3′ flanking region. The gene is transcribed by RNA polymerase III but the sequence elements both within the gene for H1 RNA and its 5′ flanking region are characteristic not only of genes transcribed by polymerase III but also of genes transcribed by polymerase II. That sequences other than those characteristic of only polymerase III transcription may be required for maximum efficiency of transcription is indicated by that fact that the gene for H1 RNA, unlike the gene for 5S rRNA, is transcribed *in vitro* much more efficiently in whole cell extracts than in the S100 and thus appears to require factors other than those required in the transcription of 5S rRNA.

The sequence beyond the 3′ terminal TTTTTGG in the gene for H1 RNA is a poly A sequence similar to those sometimes found at the ends of pseudogenes (22). However, this gene is not truncated in comparison to the cDNA sequence nor does it contain any deletions nor is it surrounded by direct repeats. Additionally, four isolates of the gene from the original genomic library appeared to carry the gene on identical restriction fragments and sequence analysis of these isolates showed that the 3′ terminal run of A residues was present in each. These facts, and most importantly, the data from the transcription experiments *in vitro* indicate that the gene we isolated is not a pseudogene.

Beyond the run of A residues at the 3′ end of the gene for H1 RNA lies an Alu element. We do not know if this Alu element has any role in transcription. However, there are 66 nucleotides deleted from the 3′ end of the Alu element in pMBH2 and this deletion does not affect transcription. We note that the genes for 7SK have Alu sequences immediately downstream from their 3′ ends (6). Furthermore, the gene for mouse MRP RNA contains a B1 element downstream from its 3′ end. The B1 element is found in rodents and is analogous to Alu elements in human DNA (8). The MRP RNA, part of an enzymatically active ribonucleoprotein particle, may be related to H1 RNA in humans (23). Both the 7SK and the MRP RNA genes resemble the H1 RNA gene in the sequences of their transcriptional control elements.

The array of potential transcriptional control elements present in the gene for H1 RNA resembles those found in a class of genes transcribed by RNA polymerase III, of which U6 is the best characterized representative (12,14,18,19,24). These genes have an internal box A and a run of T residues at their 3′ ends as required for transcription by RNA polymerase III. However, transcription of some of these genes, (the 7SK and the U6 genes), by RNA polymerase III has been found not to require the box A sequence but rather to depend upon the presence of the several sequences in the 5′ flanking region, which are most often utilized in transcription by RNA polymerase II. The run of T residues

at the 3′ end of the coding region of these genes is required for proper transcription and termination (12,17−19).

Exactly which of the various sequence elements present within and flanking the H1 RNA gene are required for efficient transcription remains to be seen. We note that in the 5′ flanking region the TATA box and the proximal sequence elements are located at almost precisely the same distances from the site of initiation of transcription (presumed to be defined by the first nucleotide of the H1 RNA; 4) as are these same sequence elements in the U6 gene (14). The octamer sequence, however, is found in the H1 RNA gene about 100 nucleotides upstream from the site of initiation of transcription, while it is located over 200 bases from this site in the U6 genes and the genes for other U RNAs (14). It is also found in this more distal location in the 7SK and Mouse RNase MRP RNA genes (8,17). Although its position is not in accord with its position in these other genes, the octamer sequence in the H1 RNA gene matches the consensus sequence perfectly. It also lies in close proximity to a Sp1 element and it has been shown that the octamer sequence and the Sp1 sequence can act together as enhancers of transcription (25). Nevertheless, in the gene for H1 RNA, deletion of all of the Sp1 elements, including the one closest to the octamer, does not affect the efficiency of transcription *in vitro*. Whether the octamer sequence itself influences transcription *in vitro* remains to be seen.

It has been suggested that genes which combine elements of both polymerase II and polymerase III promoters may be remnants of primordial eukaryotic genes (17,24). It may be that the gene for the eukaryotic version of the RNA subunit of RNase P is also very ancient, having originated at the time of divergence of eukaryotes from prokaryotes. In accord with this hypothesis is the suggestion that the catalytic RNA subunit of prokaryotic RNase P (26,27) first appeared in a time when RNA molecules performed all coding and enzymatic functions in living organisms.

ACKNOWLEDGEMENTS

We thank Donna Wesolowski and Steven Bird for excellent technical assistance and Dr. George McCorkle for helpful discussions regarding sequence identities and comparisons. We also thank Dr. D. Pravtcheva and Ms. Lisa Hunihan, of the laboratory of Professor Frank Ruddle, and Ms. Joan Menninger, of the laboratory of Professor David Ward, for performing the chromosome mapping experiments. We are grateful to Pat Maroney and Greg Hannon for performing transcription analyses *in vitro*. This work was supported by grants NIH GM-19422 and NSF DMB-8722644 to S.A. and GM-31528 to T.W.N.

REFERENCES

1. Altman,S., Gold,H.A. and Bartkiewicz,M. (1988) In Birnstiel,M.L. (ed.), Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles. Springer Verlag, Berlin, New York, pp.183−195.
2. Baer,M., Lumelsky,N., Guerrier-Takada,C., and Altman,S. (1989) In Eckstein,F., and Lilley,D.M. (ed.), Nucleic Acids and Molecular Biology. Springer Verlag, Berlin, New York, Vol 3, pp. 231 −250.
3. Altman, S. (1989) In Meister,A. (ed.), Advances in Enzymology. John Wiley and Sons Inc., pp. 1−36.
4. Bartkiewicz, M., Gold,H., and Altman,S. (1989) Genes and Development, **3**, 488−499.
5. Dahlberg,J.E., and Lund,E. (1988) In Birnstiel,M.L. (ed.), Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles. Springer Verlag , Berlin, New York, pp.38−70.
6. Murphy,S.M., Tripodi, and Melli,M. (1986) Nucl. Acids Res. **14**, 9243−9260.
7. Ullu,E., and Weiner,A. (1985) Nature, **318**, 371−374.
8. Chang,D.D., and Clayton,D.A. (1989) Cell, **56**, 131−139.
9. Howe,J.G., and Shu,M.D. (1989) Cell, **57**, 825−834.
10. Ciliberto,G., Raugei,G., Costanzo,F., Dente,L., and Cortese,R. (1983) Cell, **32**, 725−733.
11. Bogenhagen,D.F., and Brown,D.D. (1981) Cell **24**, 261−270.
12. Carbon,P., Murgo,S., Ebel,J.P., Krol,A., Tebb,G., and Mattaj,I.W. (1987) Cell, **51**, 71−79.
13. Jones,N.C., Rigby,P.W.J., and Ziff,E.B. (1988) Genes and Development, **2**, 267−281.
14. Krol,A., Carbon,P., Ebel,J.P., and Appel, B. (1987) Nucl. Acids Res.,**15**, 2463−2478.
15. Lobo,S.M., and Hernandez,N. (1989) Cell, **58**, 55−67.
16. Kadonaga,J.T., Jones,K.A., and Tijan,R. (1986) TIBS, **11**, 20−23.
17. Murphy,S., Di Liegro,C., and Melli,M. (1987) Cell, **51**, 81−87.
18. Das,G.D., Henning,D., Wright,D., and Reddy,R. (1988) EMBO J., **7**, 503−512.
19. Kunkel,G.R., and Pederson,T. (1988) Genes and Development, **2**, 196−204.
20. Quentin,Y. (1988) Journal of Molecular Evolution, **27**, 194−202.
21. Bently,K.L., Ferguson-Smith,A., and Ruddle,F. (1988) Cancer Surveys, **7**, 267−294.
22. Weiner,A.M., Deininger,P.L., and Efstratiadis,A. (1986) Ann. Rev Biochem., **55**, 631−661.
23. Gold,H.A., Topper,J.N., Clayton,D.A., and Craft,J. (1989) Science, in press.
24. Mattaj,I.W., Dathan, N.A., Parry, H.W, Carbon, P., and Krol,A. (1988) Cell, **55**, 435−442.
25. Ares M.,Jr., Chung,J.S., Giglio,L., and Weiner,A.M. (1987) Genes and Development, **1**, 808−817.
26. Altman, S. (1984) Cell, **36**, 237−239.
27. Weiner,A., and Maizels,N. (1987) Proc. Natl. Acad. Sci. USA, **84**, 7383−7387.
28. Southern, E. (1975) J. Mol. Biol., **98**, 503−517.
29. Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
30. Kaslow,D.C. (1986) Nucl. Acids Res., **14**, 6767.
31. Yanisch-Perron,C., Vieira,J., and Messing,J. (1985) Gene, **33**,103−119.
32. Weil,P.A., Segall,J., Harris,B., Ng,S.Y., and Roeder, R.G. (1979) J. Biol. Chem., **254**, 6163−6173.
33. Dignam,J.D., Lebovitz,R.M, and Roeder, R.G. (1983) Nucl Acids Res., **11**, 1475−1489.
34. Vournakis,J., Efstratiadis,A., and Kafatos, F. (1975) Proc. Natl. Acad. Sci.USA, **72**, 2959−2963.
35. Bogenhagen,D.F., Sakonju,S., and Brown,D.D. (1980) Cell, **19**, 27−35.