

## Expression of histone-U1 snRNA chimeric genes: U1 promoters are compatible with histone 3' end formation

Duane R. Pilch and William F. Marzluff

Department of Chemistry, Florida State University, Tallahassee, Florida

Chimeric genes which fuse the mouse histone H2a gene and the mouse U1b gene were constructed and introduced into CHO cells by cotransfection. In the UH genes, the U1b gene promoter and the start of the U1b gene were fused to the H2a gene in the 5' untranslated region. In the HU genes, the U1b 3' end was inserted into the 3' untranslated region of the H2a gene replacing the normal histone 3' end. Transcripts from the UH genes initiated at the start of the U1 gene and ended at the normal histone 3' end. Transcripts from the HU chimeric genes did not end at the U1 3' end but extended at least 80 nucleotides further and had heterogeneous 3' ends. Placing both a U1 snRNA promoter and a U1 snRNA 3' end around a histone coding region resulted in transcripts which initiate and terminate at the appropriate U1 ends. These results are consistent with previous reports that formation of the U1 3' ends require U1 promoters, but indicate that the histone 3' end can be formed on transcripts initiating at U1 promoters. The transcripts initiated at the U1 start site and ending at the histone 3' end are present on polyribosomes and show proper posttranscriptional regulation.

RNA polymerase II transcribes a variety of genes, including the polyadenylated mRNAs, the non-polyadenylated histone mRNAs and the U series of small nuclear RNAs which contain 5' trimethylated cap structures. These classes of transcripts have different 3' ends. The polyadenylated mRNAs are formed by an endonucleolytic cleavage 15–30 nucleotides 3' to an AAUAAA sequence, followed by polyadenylation (Birnstiel et al., 1985). A less highly conserved GU rich sequence located 3' of the cleavage site is also required for efficient processing (Proudfoot, 1989). The histone 3' ends are also formed by an endonucleolytic cleavage requiring a bipartite recognition sequence: a highly conserved stem-loop just 5' to the cleavage site

and a less highly conserved purine-rich sequence 3' to the cleavage site, which binds the U7 snRNP (Birchmeier et al., 1984; Gick et al., 1986; Mowry and Steitz, 1987).

In contrast, formation of the 3' ends of the snRNAs transcribed by RNA polymerase II requires that the polymerase initiate from an snRNA promoter (Hernandez and Weiner, 1986; Neuman de Vegvar et al., 1986). There is a relatively poorly conserved consensus sequence 3' to the end of the primary transcript (Ach and Weiner, 1987) and no absolute requirement for sequences in the transcript for proper 3' end formation (Hernandez and Weiner, 1986). If a herpes virus thymidine kinase, globin, or adenovirus promoter replaces the snRNA promoter,

Received November 13, 1990; revision accepted December 19, 1990.

Correspondence: William F. Marzluff, Department of Chemistry, Florida State University, Tallahassee, FL 32306 Tel (904)644-5282 Fax (904)644-8281

© 1991 by the University of Health Sciences/The Chicago Medical School. All rights reserved. 1052-2166/91/0101/041-13\$2.00

then the snRNA 3' end is not formed (Hernandez and Weiner, 1986; Neuman de Vegvar et al., 1986), transcription continues past the normal site, and the transcripts are polyadenylated using cryptic or natural polyadenylation sites. In addition, longer read-through transcripts formed in isolated nuclei are not precursors to mature U1 RNA molecules (Lobo and Marzluff, 1987). Potential precursors to U1 snRNA are not processed when they are injected into *Xenopus* oocytes (Neuman de Vegvar et al., 1986; Dahlberg and Lund, 1989). These results suggest that a sequence in the snRNA promoter is required for proper 3' end formation which occurs co-transcriptionally, either as a termination event or a very rapid processing event.

There is also a report that transcripts which initiate at snRNA promoters do not utilize polyadenylation sites and do not form functional mRNAs (Dahlberg and Schenborn, 1988). Taken together, these results indicate that the synthetic machinery for the production of the polyadenylated mRNAs and snRNAs are separate. Since histone mRNAs, like U snRNAs, lack introns and are not polyadenylated, we have examined the expression of chimeric histone-snRNA transcripts. The snRNA promoter will direct formation of functional histone mRNAs. However, the histone mRNA promoter directs formation of transcripts which are not terminated/processed at the snRNA 3' end but end heterogeneously 3' of the normal 3' end. However, if both the U1 promoter and U1 3' end are attached to a histone coding region, defined transcripts initiating at the U1 start site and ending at the U1 3' end are formed.

## Materials and methods

### Construction of cloned genes

The chimeric U1 and histone H2a genes were constructed from the mouse histone H2a.614 gene (Graves et al., 1985; Hurt et al., 1989) and the mouse U1b.1 and U1b.2 genes (Marzluff et al., 1983; Moussa et al., 1987). A U1 deletion mutant with a minimal promoter containing 226 nucleotides of 5' flanking sequence was digested with HphI (which cuts at nucleotide 50) or BstNI (which cuts at nucleotide 5) in the coding region, the ends blunted with the Klenow fragment of DNA polymerase and the fragment cloned into pUC18 to give the promoter cas-

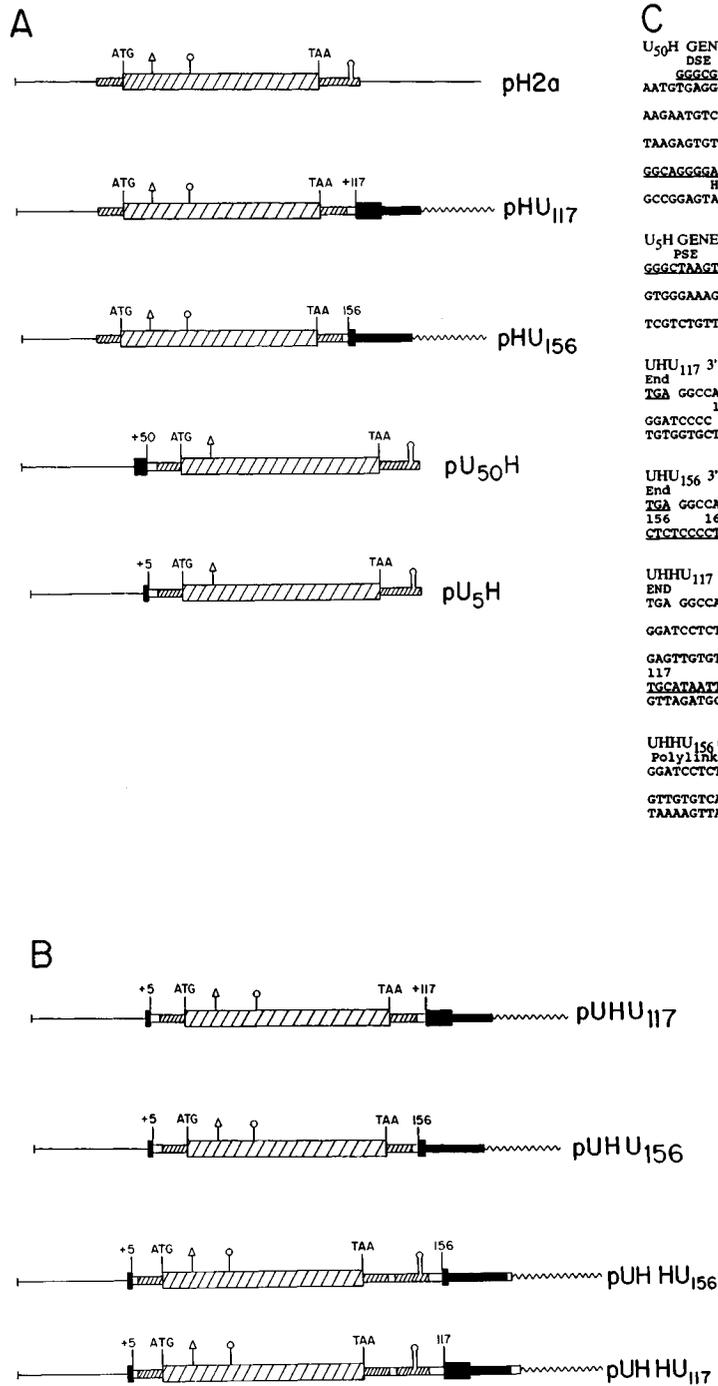
ettes, pU1<sub>5</sub> and pU1<sub>50</sub>. Two U1 3' ends with different amounts of coding region were cloned by digesting the U1b.2 gene at nucleotide 156 with CfoI, blunting the end with T4 polymerase and releasing the 3' end fragment with BglII. This fragment was inserted into pUC19 digested with BamHI and HincII to give pU1<sub>156</sub>. The U1b.1 gene was digested with TaqI at nucleotide 117 and the 130 nucleotide EcoRI to TaqI fragment was inserted into pUC18 digested with EcoRI and AccI to give pU1<sub>117</sub>. The H2a.614 coding region plus the 3' end was cloned by digestion of the EcoRI to XbaI fragment extending from -812 to 320 nucleotides past the end of the stem-loop with HphI which cuts at nucleotide -9 between the TATAA box and normal transcription start site. The end was blunted with the Klenow fragment of DNA polymerase, the histone coding region plus the 3' end released by digestion with XbaI, and the fragment inserted into either pU1<sub>5</sub> or pU1<sub>50</sub> to give pU1<sub>5</sub>H and pU1<sub>50</sub>H respectively. The H2a.614 promoter plus coding region was cloned as an EcoRI-SstI subclone which ends 20 nucleotides before the 3' stem-loop. This fragment was inserted into the pU1<sub>156</sub> and pU1<sub>117</sub> constructs to give pHU<sub>117</sub> and pHU<sub>156</sub>. The histone promoter was replaced by the U1 promoter to make the UHU<sub>156</sub> and UHU<sub>117</sub> genes. The UHHU genes were constructed by adding either the U<sub>117</sub> or U<sub>156</sub> 3' end onto a complete histone gene. The structure of the genes is shown in Figures 1A and 1B, and the sequence across the junctions is shown in Figure 1C.

### Transfection

The genes were introduced into CHO cells by cotransfection with the pSVneo gene using the polybrene procedure of Chaney et al. (1986). Stable transfectants were isolated by selection with G418 as previously described (Levine et al., 1988; Liu et al., 1989). Pools of transfectants, 20-50 per flask, were pooled and grown in the absence of G418 for analysis of expression of the transfected genes. One transfection, CHO-156C, yielded very few (<5) colonies.

### Preparation and analysis of RNA

Total cell RNA was prepared from exponentially growing cells (<50% confluent) as previously described (Levine et al., 1988). In some experiments cells were fractionated into nuclear and



cytoplasmic fractions. RNA was analyzed by S1 nuclease mapping using the genes 3' labeled at the NarI site (at codon 45) or the AvaI site (at codon 20) of the H2a gene to map the 3' end and labeled at the 5' end of the NarI site or the AvaI site to map the 5' end. The conditions of hybridization and digestion have been

C

**U<sub>50</sub>H GENES**  
 DSE  
GGGCTGGTGTATGCAGATAAGCGCGCAGGGTCGGCGGGTCCGTCTCCGTGCAGA  
 AATGTGAGGGGTTTCTCCCGCGGAGGGAGGGAGCGCGGATACGGTCTAGCAGAGCGC  
 PSE  
 AAGAATGTCGTGCGGACGAGGGAGACCGAGATCGGCCAGGGCTAAGTACCGTGTGT  
 +1 U1  
 TAAGAGTGTAGCGCGGACGGGTGTGAGCCGGGCGCGGACAGTGGGAAAGCATACTTACCT  
 Polylinker  
GGCAGGGGAGATACCATGATCATGAAGGTGGTTTTCCAG GACTCTAGAGGATCCCC  
 Histone sequence  
 GCGGAGTACCGTTCTCGTGTGGCCGTTCGTCTGTTCGGCTTTCGTGCATGTCCGGT

**U<sub>5</sub>H GENES (Starting at proximal sequence element)**  
 PSE U1 Sequence  
GGCTAAGTACCGTGTGTTAAGAGTGTAGCGCGGACGGGTGTGAGCCGGGCGGGCA  
 +1 Polylinker Histone sequence  
 GTGGGAAAGCATACT CTAGAGGATCCC CCGCGGAGTACCGTTCTCGTGTTCGGCT  
 HetSerGly  
 TCGTCTGTTCGGCTTTCGTGCATGTCCGGT

**UHU<sub>117</sub> 3' END**  
 End Histone sequence Polylinker  
 TGA GGCCACTGCTCGCGTCGGCCCGGCTCTCTGTGATATAACCCCGCGAGCTC GGTACCCGG  
 117 U1 coding sequences 165  
 GGATCCCC GGATGCAATTTCTGGTACTGGGGACTGGGTTCCGGCTCTCCCGTATTTT  
 TGTGCTCTAAAAGTTAGATGCATTCTGCTCTCTCATGTCTCTTTACATGTGTTTGTG

**UHU<sub>156</sub> 3' END**  
 End Histone sequence Polylinker  
 TGA GGCCACTGCTCGCGTCGGCCCGGCTCTCTGTGATATAACCCCGCGAGCTC GGTACCCGG  
 156 165 U1 sequence  
CTCTCCCGTATTTTTGTGGTGTAAAAGTTAGATGCATTCTGCTCTCTCATGTCTCTTT

**UHHU<sub>117</sub>**  
 END Histone sequence Polylinker  
 TGA GGCCACTGCTCGCGTCGGCCCGGCTCTCTGTGATATAACCCCGCGAGCTC GGTACCCGG  
 Histone 3' end  
 GGATCTCTAGATTC GAGCTCCCAAAAAGGCTCTTTGAGAGCCACCCCACTGAATCAGATAAA  
 Polylinker  
 GAGTTGTGTCAAGGTAGCC GACCTGCAGGCATGCAAGCTGGGGATCCCGCGAC  
 117 U1 sequence 165  
TGCATAATTTGTGGTACTGGGGACTCGGTTCCGGCTCTCCCGTATTTTGTGGTGTAAAA  
 GTTAGATGCATTCTGCTCTCTCATGTCTCTTTACATGTGTTTGTGAGGCATGGCAGCAAT

**UHHU<sub>156</sub> (starting 60 nts 3' to histone coding region)**  
 Polylinker Histone 3' end  
 GGATCTCTAGATTC GAGCTCCCAAAAAGGCTCTTTGAGAGCCACCCCACTGAATCAGATAAA  
 Polylinker 156 165 U1 sequence  
 GTTGTGTCAAGGTAGCC GACCTGCAGGCATGCAAGCTGGGGA CTCTCCCGTATTTTTGTGGT  
 TAAAAGTTAGATGCATTCTGCTCTCTCATGTCTCTTTACATGTGTTTGTGAGGCATGGCTCGGA

**Figure 1.** The histone-U1 chimeric genes. **A.** The parent histone H2a-614 gene is shown (pH2a), with the 5' and 3' untranslated regions and the coding region indicated. The U<sub>5</sub>H and U<sub>50</sub>H gene have 226 nucleotides of 5' flanking sequence and 5 or 50 nucleotides of the mouse U1b coding region. The HU<sub>117</sub> and HU<sub>156</sub> genes have the U1 3' end starting at nucleotides 156 and 117 respectively, attached in the 3' untranslated region of the histone H2a<sub>614</sub> gene. **B.** The UHU<sub>156</sub> and UHU<sub>117</sub> genes have both the U1 promoter and U1 3' end flanking the histone gene. The UHHU genes have the U1 3' end starting at nucleotide 156 or 117 of the coding region attached 25 nucleotides 3' of the end of the U<sub>5</sub>H gene. **C.** The sequences of the U1-histone junctions in the U<sub>5</sub>H, U<sub>50</sub>H, HU<sub>117</sub>, HU<sub>156</sub> genes are given. The coding region sequences of the U1 genes are underlined, as are the distal sequence element (DSE), proximal sequence element (PSE) of the U1 promoter, the ATG codon of the H2a-614 gene, and the 3' end of the histone mRNA. The polylinker sequences are indicated. The restriction sites used are: † AvaI; ‡ NarI.

described (Graves et al., 1985), and the temperature of hybridization was 58°C for all probes. The protected fragments were resolved by electrophoresis on 6% polyacrylamide-7M urea gels and detected by autoradiography.

## Results

In order to test the compatibility of histone promoters and 3' ends with the snRNA promoters and 3' ends, we constructed the four chimeric genes with a U1 promoter or U1 3' end replacing the histone promoter or 3' end shown in Figure 1A. These genes all contain the intact mouse histone H2a-614 coding region, with either the U1 promoter or the U1 3' end replacing the corresponding histone sequence. The U1 promoter sequence included either 5 or 50 nucleotides of the U1 coding region and was attached 9 nucleotides prior to the normal start of transcription of the H2a-614 gene. The U1 3' end including the last 49 or 10 nucleotides of the coding region was attached in the H2a-614 3' untranslated region in place of the normal stem-loop sequence.

In Figure 1B we show additional constructs which contain U1 promoters and U1 3' ends. Two constructs (UHU<sub>117</sub> and UHU<sub>156</sub>) had the U1<sub>5</sub> promoter and two different U1 3' ends flanking the complete histone coding region. Finally in the UHHU<sub>117</sub> and UHHU<sub>156</sub> genes, the U1<sub>5</sub> promoter was followed by a histone H2a coding region and H2a 3' end and then the U1 3' end was placed 3' of the histone 3' end. The sequences at the junctions of the U1 and histone genes in the four chimeric transcripts are shown in Figure 1C.

These eight genes were introduced into CHO cells in low copy number by cotransfection with the neomycin resistance gene and stable transformants selected for resistance to G418. To analyze for expression of the genes we utilized S1 nuclease assays, which allowed us to measure the 5' and 3' ends of the transcripts.

### Transcripts which initiate at U1 promoters form histone 3' ends

Figure 2 shows the expression of the U1-histone (UH) chimeric genes. Transfection of CHO cells by this procedure results in low copy number transfections which show levels of histone gene expression which reflect the activity of the genes in mouse cells (Levine et al., 1988). We have previously shown that the H2a-614 gene is expressed

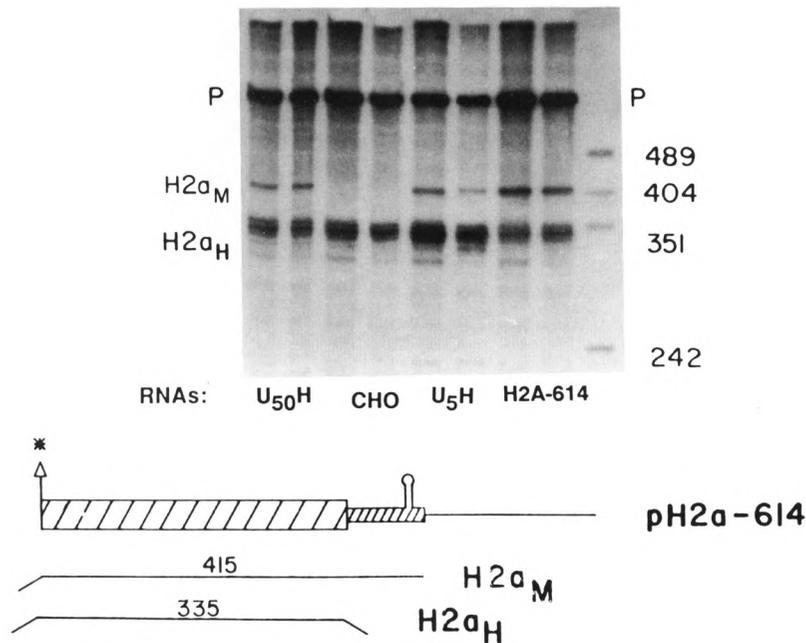
very efficiently in CHO cells (Levine et al., 1988). The 3' S1 nuclease assays show that large amounts of transcripts which end at the histone 3' end accumulated in CHO cells transfected with the U<sub>5</sub>H or U<sub>50</sub>H gene (Fig. 2A). Similar amounts of transcripts (about 50% of the amount expressed from the H2a-164 gene), accumulated in the cells transfected with the U1/H2a chimeric genes as in the cells transfected with intact H2a-614 gene (Fig. 2A), indicating that the U1 promoter efficiently directed transcription of histone mRNA.

Figure 2B shows the 5' ends of these transcripts mapped using the U<sub>50</sub>H or U<sub>5</sub>H genes labeled at the 5' end of the *Ava*I site located at codon 20 in the histone coding region. All of the transcripts from the U<sub>50</sub>H gene initiated at the U1 5' end (Fig. 2B, lanes 1 and 2). Note that the same sample was analyzed in lanes 1 and 7. In the mapping in lane 7 all the transcripts from the U<sub>50</sub>H gene are mapped to the same site, the point of divergence from the U<sub>5</sub>H gene. Since the protected fragments in lane 1 and 7 are the same intensity relative to the endogenous histone mRNA, this indicates that there are not significant amounts of read-through transcripts which initiated upstream of the normal U1 start site. Similar results were found for the U<sub>5</sub>H gene (Fig. 2B, lanes 5 and 6). Again we can rule out the presence of read-through transcripts which originate upstream of the U1 start site by comparing the intensities of the bands in lanes 3 and 5, which are assays of the same RNA sample. RNA from cells transfected with the intact H2a-614 gene was also assayed (Fig. 2B, lanes 4 and 8), and these assays confirmed that the levels of expression of histone mRNAs from the U1 and histone promoters were similar. Since there were similar relative intensities obtained for both the 5' and 3' ends of the UH mRNAs and the H2a-614 mRNAs, we can conclude that the great majority of the transcripts which initiated at the U1 promoter went all the way to the histone 3' end.

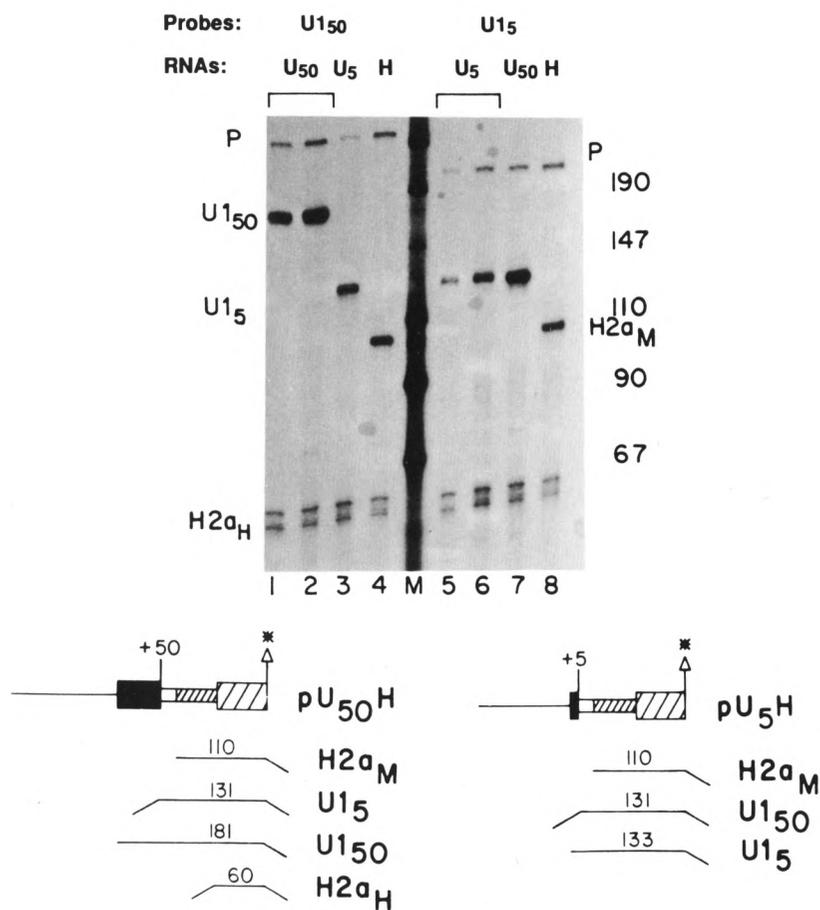
### The UH mRNAs show proper posttranscriptional regulation

The U<sub>5</sub>H RNA should encode an authentic histone protein, since the first AUG in the transcript at position 71 is the same AUG codon which initiates translation of the H2a.2 protein, and the complete coding sequence is present. To determine whether transcripts which initiated at a U1 promoter formed functional

**A**

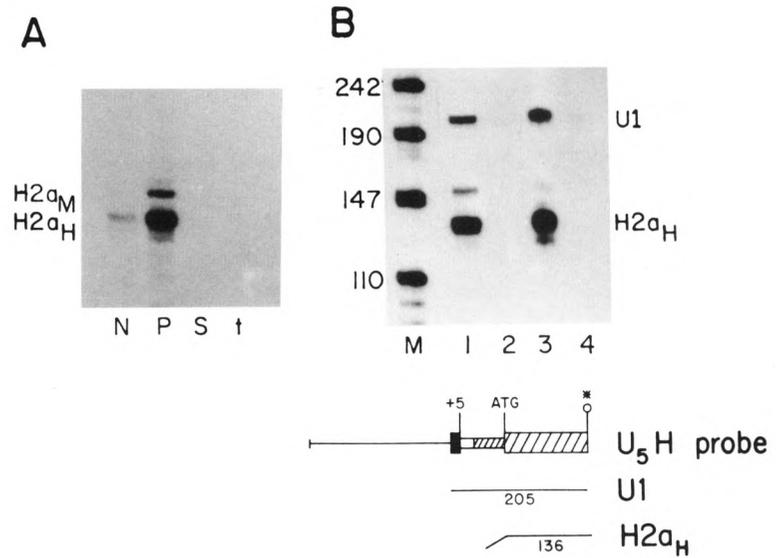


**B**

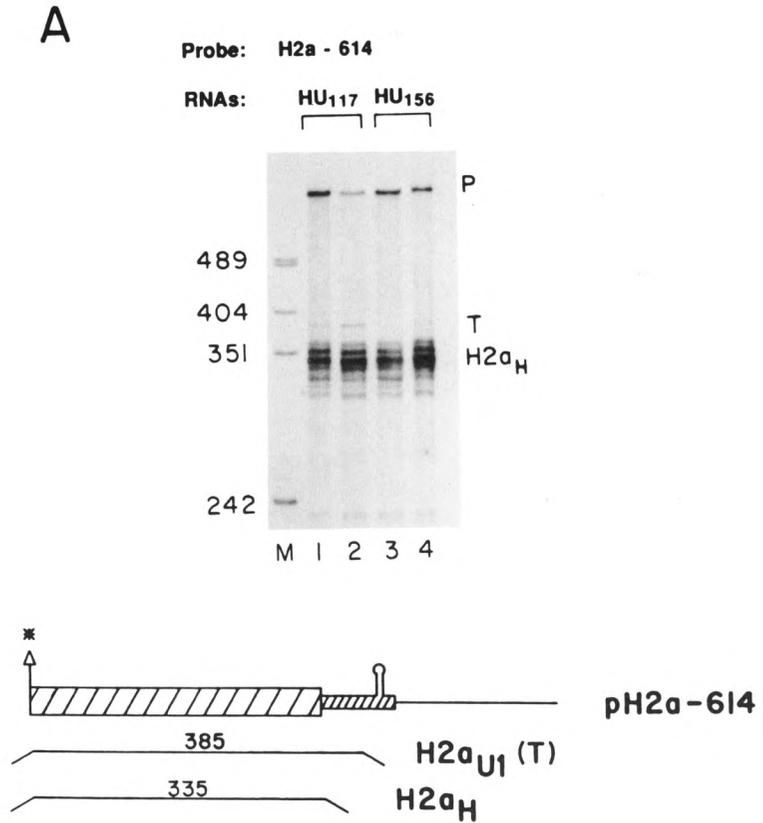


**Figure 2.** Expression of the UH genes. **A.** 5  $\mu$ g RNA from duplicate samples of CHO cells transfected with the U<sub>5</sub>H, U<sub>50</sub>H or H2a-614 genes were analyzed by S1 nuclease mapping using the H2a-614 gene labeled at the 3' end of the *Ava*I site at codon 20 as a probe. The protected fragments are: H2a<sub>H</sub>—protection of the endogenous hamster H2a mRNAs. H2a<sub>M</sub>—protection of the RNA expressed from the transfected genes. Band P is the probe band. The RNA samples are indicated under each lane. In the lanes labeled CHO, RNAs from untransfected CHO cells were analyzed. The marker (last lane on the right) is pUC18 digested with *Hpa*II. **B.** The same RNA samples as in **A** were analyzed by S1 nuclease mapping using either the U<sub>50</sub>H gene (lanes 1–4), or the U<sub>5</sub>H gene (lanes 5–8) labeled at the 5' end of the *Ava*I site as a probe. Lanes 1, 2, and 7 are analyses of RNAs from cells transfected with the U<sub>50</sub>H gene, Lanes 3, 5, 6 are analyses of RNAs from cells transfected with the U<sub>5</sub>H gene. Lanes 4 and 8 are analyses of RNAs from cells transfected with the H2a-614 gene. Lanes 2 and 7 and lanes 3 and 6 contain the same RNA samples. The protected fragments are: H2a<sub>H</sub>—protection of the endogenous hamster H2a mRNAs; H2a<sub>M</sub>—protection of the mouse H2a-614 mRNA; U1<sub>5</sub>—protection of the RNA from the U<sub>5</sub>H gene starting at the U1 RNA start site; U1<sub>50</sub>—protection of RNA from the U<sub>50</sub>H gene starting at the U1 start site. P indicates the undigested probe fragment. In lanes 3 and 7, the RNAs are mapped to the point of divergence between the U1<sub>5</sub>H and U1<sub>50</sub>H genes. Lane M is marker pUC18 digested with *Hpa*II.

**Figure 3. A.** Localization of the U<sub>5</sub>H transcripts. CHO cells transfected with the U<sub>5</sub>H gene were fractionated into nuclear (N), polysomal (P) and cytosol (S) fractions and equal cell equivalents of each fraction assayed by S1 nuclease mapping using the H2a-614 gene labeled at the 3' end of the Aval site at codon 20 as a probe as in Figure 2A. The protected fragments are: H2a<sub>H</sub>—protection of the endogenous hamster H2a mRNAs; H2a<sub>M</sub>—protection of the RNA expressed from the U<sub>5</sub>H gene. **B.** Effect of hydroxyurea on U<sub>5</sub>H RNA. Duplicate cultures of CHO cells transfected with the U<sub>5</sub>H gene were treated with 5mM hydroxyurea for one hour and RNA prepared from control (lanes 1 and 3) and treated cultures (lanes 2 and 4). 5 μg total cell RNA was assayed by S1 nuclease mapping using the U<sub>5</sub>H gene labeled at the 5' end of the NarI site at codon 43 as a probe. Lane M is marker pUC18 digested with HpaII. The protected fragments are: H2a<sub>H</sub>—protection to the ATG codon by the endogenous hamster H2a mRNAs; U1—protection to the first nucleotide in the U1 sequence of the U<sub>5</sub>H gene.



**Figure 4. Expression of the HU genes.** **A.** Duplicate preparations of RNA from cells transfected with the HU<sub>117</sub> genes (lanes 1 and 2) or the HU<sub>156</sub> genes (lanes 3 and 4) were analyzed by S1 nuclease mapping using the H2a-614 gene labeled at the 3' end of the Aval site as a probe. This probe will map all transcripts from the HU genes which extend more than 60 nucleotides past the 3' end of the histone coding region as a single protected fragment (labeled T). The other protected fragment (H2a<sub>H</sub>) is due to protection of the endogenous hamster H2a mRNAs; **B.** Duplicate preparations of RNA (5 μg) from cells transfected with the HU<sub>117</sub> (lanes 1 and 2) or the HU<sub>156</sub> genes (lanes 5 and 6) were analyzed by S1 nuclease mapping using the HU<sub>117</sub> (lanes 1-4) or HU<sub>156</sub> (lanes 5-8) genes labeled at the 3' end of the NarI site at codon 43 as probes. RNA from cells transfected with the H2a-614 gene was analyzed in lanes 3, 4, 7, and 8. The arrows indicate the expected positions of transcripts ending at the U1 3' end (332 and 376 nucleotides respectively). Band P is undigested probe. **C.** RNAs from cells transfected with the HU<sub>156</sub> or the HU<sub>117</sub> were analyzed by S1 nuclease mapping using the probes UHU<sub>156</sub>PL or UHU<sub>117</sub>PL respectively. The RNAs used are indicated beneath the lanes. These probes have a polylinker insert either 112 or 82 nucleotides after the U1 3' end, which allows the mapping of longer heterogenous read-through transcripts from these genes as a single protected fragment (labeled RT). The position of the protected fragment which would correspond to the U1 3' end is indicated. The protected fragment labeled H2a<sub>H</sub> corresponds to protection of the endogenous hamster H2a mRNAs. Lane M is marker pUC18 digested with HpaII. P is the position of the undigested probe.



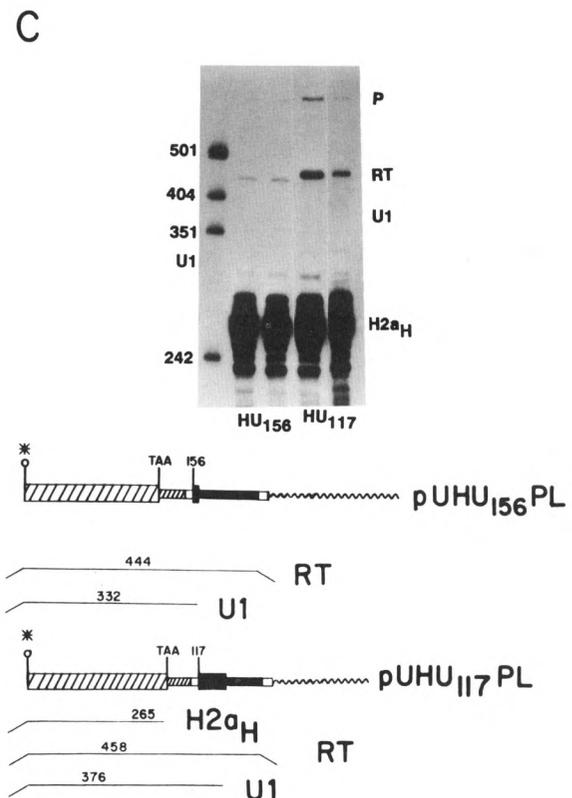
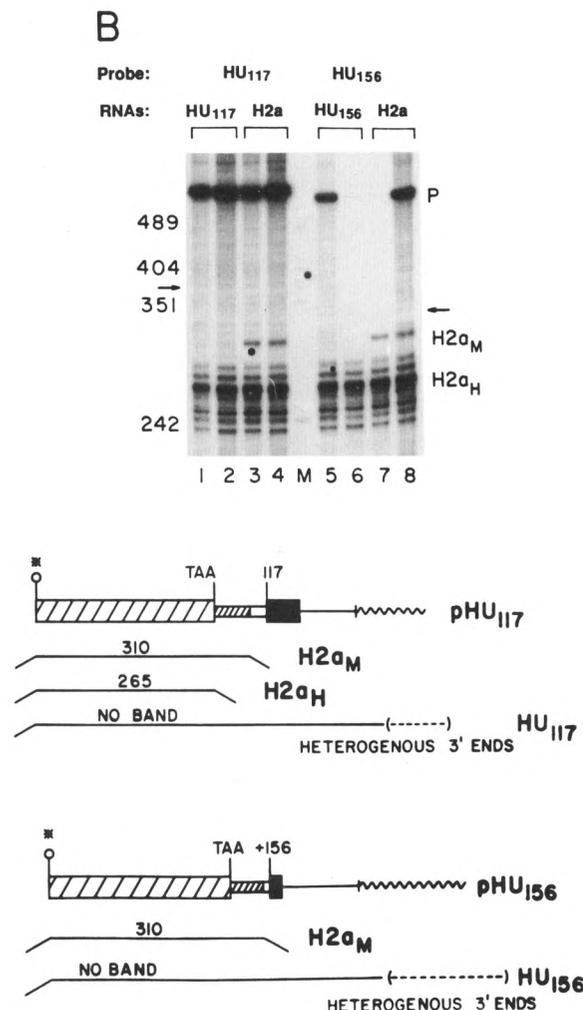
mRNAs, the cells were fractionated into nuclear (Fig. 3A, lane N), polysomal (lane P) and cytosolic (lane S) fractions. The U<sub>5</sub>H RNA fractionated in parallel with the endogenous histone mRNA and was predominantly present on polyribosomes, indicating that it is actively translated. In Figure 3A the U<sub>5</sub>H RNAs were mapped with the same 3' labeled probe used in Figure 2A. The amount of U1 RNA in each fraction was monitored and over 90% of the U1 RNA was in the nuclear fraction (not shown).

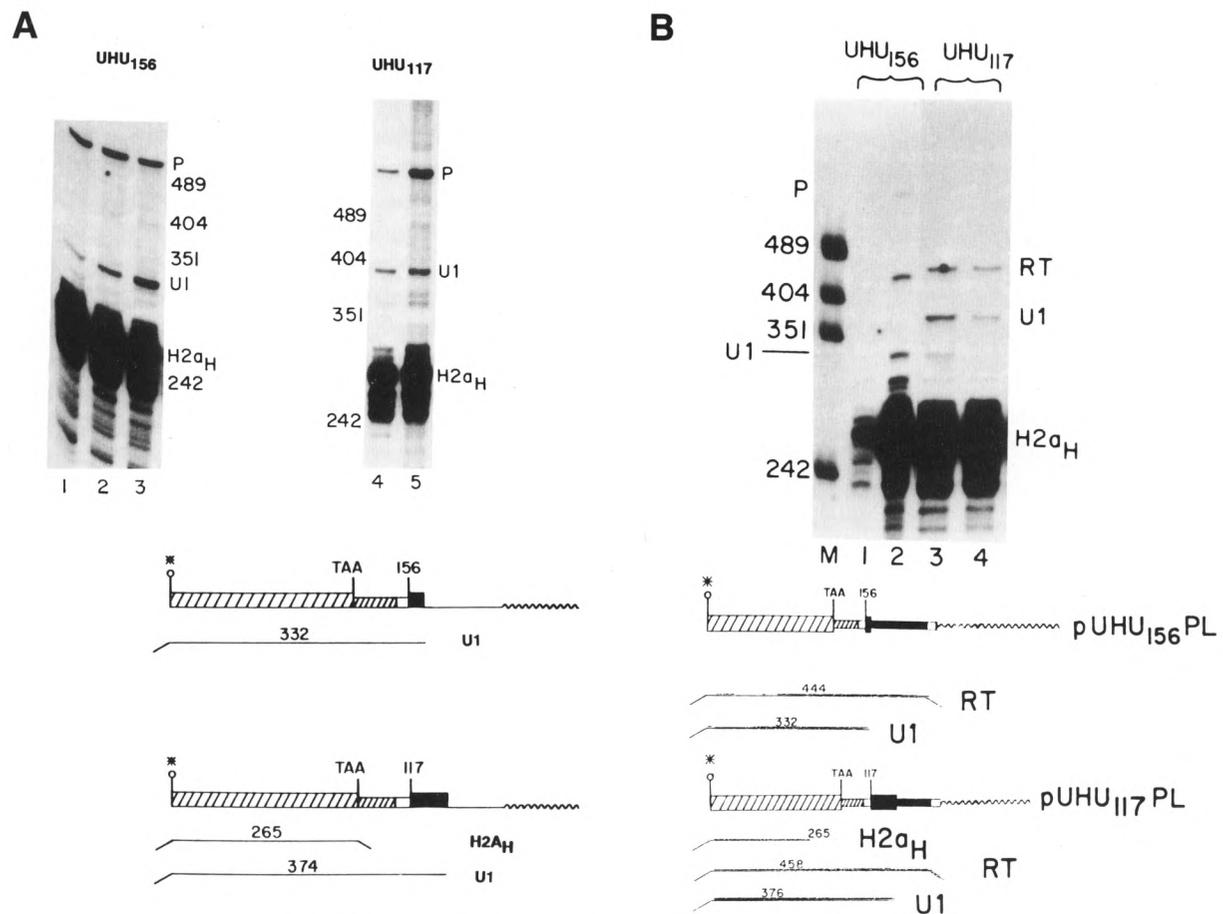
One characteristic of the regulation of histone mRNAs is that the polysomal histone mRNAs are rapidly degraded when DNA synthesis is inhibited (Sittman et al., 1983; Graves et al., 1987). Figure 3B shows that the U<sub>5</sub>H RNA shows normal posttranscriptional regulation of histone mRNA degradation which is found for all histone genes (Marzluff and Pandey, 1988). To test whether the UH mRNAs were regulated coordinately with DNA synthesis, the

transfected cells were treated with hydroxyurea and the UH RNAs mapped with a probe labeled at the 5' end of the NarI site at codon 45 in the histone H2a gene. This probe maps the endogenous hamster histone mRNAs more efficiently than the probe labeled at codon 20 (cf. Fig. 3B with Fig. 2B). The UH mRNAs and the endogenous hamster mRNAs were rapidly degraded (Fig. 3B, lanes 2 and 4). This result indicates that the UH chimeric histone mRNAs show the normal posttranscriptional regulation. This result is consistent with the finding that the regulation of histone mRNA degradation requires only the stem-loop at the 3' end of the histone mRNA (Graves et al., 1987; Pandey and Marzluff, 1987).

**Transcripts initiating at a histone promoter do not form U1 3' ends**

Figure 4 shows the S1 mapping results with the histone/U1 (HU) chimeric genes. To show that

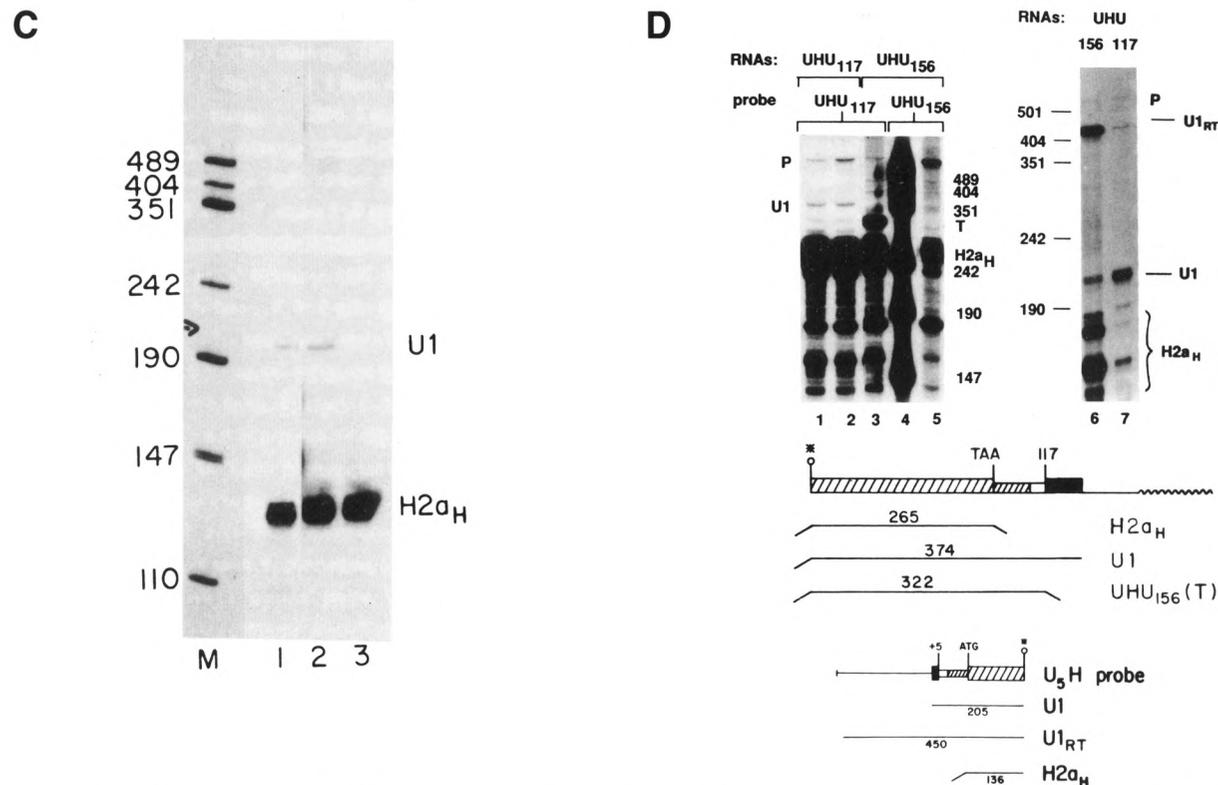




**Figure 5.** Expression of the UHU genes. **A.** 10  $\mu$ g RNA from CHO cells transfected with the UHU<sub>156</sub> genes (lanes 1–3) or the UHU<sub>117</sub> genes (lanes 4 and 5) was analyzed by S1 nuclease mapping using the appropriate homologous gene labeled at the 3' end of the NarI site at codon 43 as a probe. The protected fragments are: H2a<sub>H</sub>—protection to the TAA codon by the endogenous hamster H2a mRNAs; U1—protection by transcripts from the UHU genes extending to the normal U1 3' end. P is the undigested probe band. The positions of the marker fragments, pUC18 digested with HpaII, are indicated. **B.** 10  $\mu$ g total cell RNA from CHO cells transfected with the UHU<sub>156</sub> genes (lane 2) or the UHU<sub>117</sub> genes (lanes 3 and 4) was analyzed by S1 nuclease mapping using the UHU<sub>156</sub>PL (lanes 1 and 2) or UHU<sub>117</sub>PL (lanes 3 and 4) probes described in Fig. 4C labeled at the 3' end of the NarI site at codon 43. These probes diverge from the transfected gene after the U1 3' end as in Figure 4B, allowing detection of all longer transcripts as a single protected fragment (RT). The protected fragments are: H2a<sub>H</sub>—protection to the TAA codon by the endogenous hamster H2a mRNAs; U1—protection by transcripts from the UHU genes extending to the normal U1 3' end; RT—protection by transcripts which extend past the U1 3' end. Lane 1 is 1  $\mu$ g RNA from untransfected

transcripts were formed from these genes, S1 nuclease mapping was performed using the mouse histone H2a-614 gene as a probe (Fig. 4A). This probe maps all transcripts which extend at least 60 nucleotides past the end of the histone coding region as a single protected fragment, due to protection of the probe up to the point where the U1 sequence was attached. Clearly there were transcripts formed from the HU genes (Fig. 4A, band T). To attempt to map the ends of these transcripts we used an homologous probe (Fig. 4B). No major defined transcripts were detected using this probe, suggest-

ing that the 3' ends of these transcripts were heterogenous. In Figure 4C we used a probe which had additional polylinker sequences not in the transfected gene placed 80 nucleotides (UHU<sub>117</sub>) or 110 (UHU<sub>156</sub>) nucleotides 3' of the U1 3' end. These probes allow one to map all the transcripts which extend past the U1 end as a single protected fragment of about 450 nucleotides (RT in Fig. 4C). Figure 4C shows that transcripts extending at least 90 nucleotides past the U1 3' end are readily detected, demonstrating that the U1 3' end is not used by transcripts which have initiated at a histone promoter. Thus



CHO cells. **C.** 5  $\mu$ g total cell RNA, cells transfected with the UHU<sub>117</sub> gene (lane 1), the UHU<sub>156</sub> gene (lane 2), or untransfected CHO cells (lane 3) were analyzed by S1 nuclease mapping using the U<sub>5</sub>H gene labeled at the 5' end of the NarI site (codon 43) as a probe. The protected fragments are: H2a<sub>H</sub>—protection by the endogenous hamster H2a mRNAs; U1—protection by transcripts from the UHU genes extending to the normal U1 5' end. Lane M is marker pUC18 digested with HpaII. **D.** 10  $\mu$ g total cell RNA from cells transfected with the UHU<sub>117</sub> genes (lanes 1 and 2) or the CHO-156C cells transfected with the UHU<sub>156</sub> gene (lanes 3 and 4) were analyzed by S1 nuclease mapping using the UHU<sub>117</sub> gene (lanes 1–3) or the UHU<sub>156</sub> gene (lane 5) labeled at the 3' end of the NarI site (codon 43) as a probe. The protected fragments are labeled as in Figure 5B. The fragment labeled T is due to protection of all transcripts from the UHU<sub>156</sub> gene to the point of divergence of the UHU<sub>156</sub> and the UHU<sub>117</sub> genes. Band P is the undigested probe. Lane 4 is marker pUC18 digested with HpaII.

The same CHO-156C [UHU<sub>156</sub>] (lane 6) and UHU<sub>117</sub> RNAs (lane 7) were analyzed by S1 nuclease mapping using the U<sub>5</sub>H gene labeled at the 5' end of the NarI site as a probe. The protected fragments are labeled as in Figure 5C. The fragment labeled U1<sub>RT</sub> results from protection by transcripts which initiated upstream from the normal U1 start site. Band P is undigested probe DNA. The position of the marker fragments is shown. The hamster histone RNAs were not detected efficiently in lane 7 for unknown reasons.

transcripts which initiate at a histone promoter do not form U1 3' ends.

#### Placing a U1 promoter on the HU gene restores U1 3' end formation

Two additional genes, UHU<sub>117</sub> and UHU<sub>156</sub>, were constructed which had the U1 promoter and 5 nucleotides of the U1 coding region replacing the histone promoter (see Fig. 1B). These genes were transfected into CHO cells and the transcripts assayed by S1 nuclease mapping using an homologous probe. Figure 5A shows that there were transcripts formed ending at the U1 3' end from both of the UHU genes. To

estimate the efficiency of U1 3' end formation, these RNAs were also mapped using the probes used in Figure 4C, which can map any read-through transcripts which extend past the U1 3' end. Any transcripts extending past the U1 3' end will be detected as a single protected fragment, as discussed above (RT in Fig. 5B). Figure 5B shows that about 50% of the transcripts from both the UHU<sub>117</sub> and UHU<sub>156</sub> genes present in steady-state RNA had U1 3' ends and about 50% of the transcripts were at least 80 nucleotides longer than the U1 RNA and ended heterogeneously.

The great majority of these transcripts pres-

ent in steady-state RNA were initiated at the U1 start site. Figure 5C shows the results of S1 nuclease mapping of the 5' end of the transcripts from the same RNAs analyzed in Figure 5B. As was found for the UH genes, most—if not all—of the transcripts initiated at the U1 start site (Fig. 5C, lanes 1 and 2; see also Fig. 5D, lane 7). Similar results have been found in four independent transfections of each UHU gene, each transfection yielding a pool of more than 30 transformants.

Figure 5D shows an example of a transfection, CHO-156C, in which the UHU<sub>156</sub> gene was expressed in an unusual manner. In the CHO-156C transfection, there were less than 5 colonies obtained, and it is possible that all of the colonies were derived from a single transfectant. The 3' ends of the transcripts were mapped using two probes, the UHU<sub>117</sub> gene (Fig. 5D, lanes 1–3) and the homologous UHU<sub>156</sub> gene (Fig. 5D, lane 5). The UHU<sub>117</sub> probe will map the transcripts ending at the U1 end from the UHU<sub>117</sub> genes (Fig. 5D, lanes 1 and 2) and will map all transcripts from the UHU<sub>156</sub> gene which extend past the histone coding region as a single fragment of 322 nucleotides, the point where the UHU<sub>117</sub> and UHU<sub>156</sub> gene diverge (Fig. 5D, lane 3). There were large amounts of transcripts formed in the CHO-156C cells (Fig. 5D, lane 3), but very few, if any, of these transcripts ended at the U1 3' end when they were mapped with an homologous probe (Fig. 5D, lane 5), suggesting that the 3' ends are heterogeneous and/or extend past the end of the probe. When the 5' ends of these transcripts were mapped using a probe labeled at the 5' end of the *NarI* site at codon 45 of the H2a gene, most of the transcripts from the UHU<sub>156</sub> gene in these cells initiated upstream of the U1 start site (Fig. 5D, lane 6, band U1<sub>RT</sub>). In contrast, in all the other transfections at least 90% of the transcripts initiated at the U1 start site, although there were a small number of transcripts which initiated at the upstream start site (Fig. 5D, lane 7). It is likely that these read-through transcripts are a result of the transfected gene integrating 3' of a strong promoter. Thus in the CHO-156C cells, transcripts which initiated upstream of the U1 gene and traversed the entire U1 promoter and gene did not form U1 3' ends.

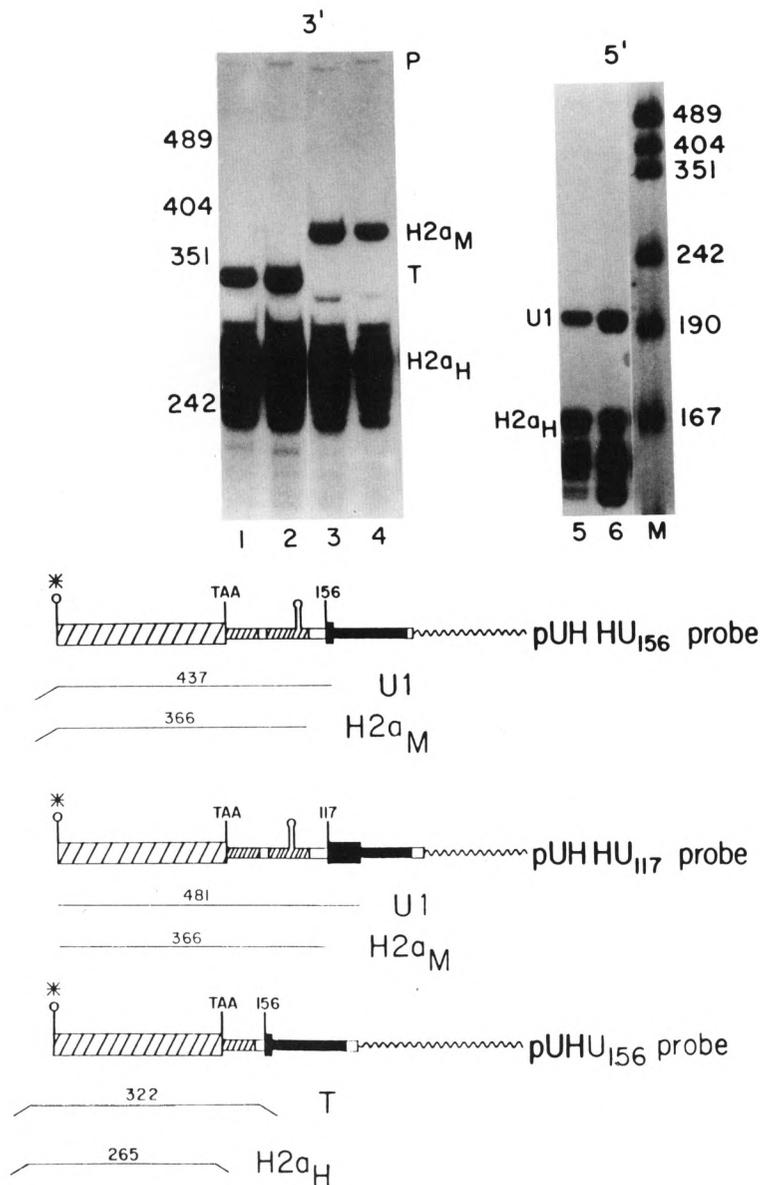
Although replacing the histone promoter and 3' end with a U1 3' end and U1 promoter on a histone gene restored formation of U1 3' ends,

the U1 3' end was not dominant to a histone 3' end. The UHHU genes (see Fig. 1B) had both a histone 3' end and U1 3' end. The 3' ends were mapped with the UHU<sub>156</sub> gene, which maps all transcripts from these genes (Fig. 6, lanes 1 and 2, band T), and with the appropriate homologous probe (Fig. 6, lanes 3 and 4). All the transcripts from these genes had histone 3' ends (Fig. 6, lanes 3 and 4). We would have detected transcripts ending at the U1 3' end if they had been present in 5% of the amount of the histone transcripts. Identical results were obtained with the UHHU<sub>156</sub> (Fig. 6, lanes 3 and 6) and UHHU<sub>117</sub> (Fig. 6, lanes 4 and 5) genes. The 5' ends of these transcripts were mapped and all the transcripts initiated at the normal U1 start site (Fig. 6, lanes 5 and 6). The presence of the histone 3' end probably did not prevent transcription continuing to the U1 3' end, since transcription normally extends several hundred nucleotides 3' of the histone genes (Chodchoy et al., 1987; Chodchoy et al., 1991). It is likely that the histone 3' ends are formed rapidly and efficiently on transcripts which initiate at a U1 promoter. However, we cannot rule out the possibility that the initial transcripts formed had U1 3' ends, which were then further processed to yield histone 3' ends.

## Discussion

The results reported here extend the previous reports that the formation of snRNA 3' ends is dependent on proper initiation from an snRNA promoter (Hernandez and Weiner, 1986; Neuman de Vegvar et al., 1986). They also demonstrate that initiation from a U1 promoter is compatible with formation of histone 3' ends, and that functional histone mRNA is produced from the UH genes. In addition, we show that transcripts initiated from a histone promoter do not form U1 3' ends, although placing a U1 promoter and U1 3' end signal around a histone gene lacking a histone 3' end results in defined transcripts which start and end at the proper U1 termini.

The coupling of the promoter to 3' end formation is a unique property of the snRNA genes. Histone genes share some common features with U snRNA genes in that they do not contain introns and do not encode polyadenylated RNAs. Transcripts which initiate from histone promoters do not form U1 3' ends; instead, the tran-



**Figure 6.** Expression of the UHHU genes. 10  $\mu$ g total cell RNA from cells transfected with the UHHU<sub>156</sub> gene (lanes 1, 4, and 5) or the UHHU<sub>117</sub> gene (lanes 2, 3, and 6) were analyzed by S1 nuclease mapping. The UHU<sub>156</sub> gene labeled at the 3' end of the NarI site as a probe in lanes 1 and 2. The UHHU<sub>117</sub> or UHHU<sub>156</sub> gene labeled at the NarI site was used as a probe in lanes 3 and 4 respectively. The 5' ends of the genes were analyzed in lanes 5 and 6 using the U<sub>5</sub>H gene labeled at the 5' end of the NarI site as a probe as in Figure 3B. Lane M is marker pUC18 digested with HpaII. The protected fragments are: H2a<sub>H</sub>—protection of the endogenous hamster H2a mRNAs; H2a<sub>M</sub>—protection by transcripts extending to the 3' end of the mouse H2a-614 gene; T—protection by all transcripts from the UHHU genes to the point of divergence between the UHHU gene and the UHU<sub>156</sub> gene; U1—protection by transcripts initiating at the U1 start site. Band P is unprotected probe DNA. A schematic of the S1 nuclease assays is shown below the figure.

scripts extend past the U1 3' end and end heterogeneously. However, histone 3' ends are formed efficiently on transcripts which initiate from a globin promoter (Whitelaw et al., 1986; Pandey et al., 1990), and polyadenylated RNAs are formed efficiently when a histone promoter drives a gene containing a polyadenylation signal (Levine et al., 1987).

In contrast, the initiation of transcription from a U1 snRNA promoter is compatible with the formation of histone 3' ends. The amount of histone mRNA formed from the UH genes is similar to the amount of RNA which accumulates from the intact histone H2a-614 gene.

Since both of these promoters are highly active promoters, it is likely that most, if not all, of the transcripts which initiate at the U1 promoter form histone 3' ends. The U<sub>5</sub>H gene also gives rise to a functional histone mRNA which is present on polyribosomes. The first AUG codon in this transcript is the AUG codon in the H2a gene, and the rest of the sequence is identical to the H2a-614 mRNA. The U<sub>5</sub>H RNA is rapidly degraded when DNA synthesis is inhibited, as is the endogenous histone mRNA. This is expected since the 3' end of histone mRNA is responsible for the posttranscriptional regulation of histone mRNA levels (Lüscher and

Schümperli, 1987; Graves et al., 1987; Pandey and Marzluff, 1987).

These results are in contrast to the results of Dahlberg and Schenborn, who did not observe any functional polyadenylated mRNA from a chloramphenicol acetyltransferase (CAT) gene driven by a U1 promoter (Dahlberg and Schenborn, 1988). This was probably at least partially due to premature termination of the transcript due to presence of cryptic 3' end signals in the CAT gene, which "terminated" transcripts which initiated at the U1 promoter (Dahlberg and Schenborn, 1988). The signals required for 3' end formation on transcripts initiated at a U1 promoter are rather flexible (Ach and Weiner, 1987). The histone H2a-614 gene is over 70% GC in the coding region (Hurt et al., 1989) and does not contain any cryptic "termination" sites. Transcripts initiated at the U1 promoter efficiently transcribe the complete histone coding region and form large amounts of histone mRNA.

There is also a report that transcripts which initiate at a U1 promoter do not efficiently utilize 3' polyadenylation sites (Neuman de Vegvar et al., 1986), although other workers have shown that in some instances transcripts initiating at U1 snRNA promoters could be polyadenylated (Hernandez and Weiner, 1986). These results differ from our results, which show that the histone 3' processing signals are efficiently used on transcripts which initiate from a U1 promoter. It is clear from these results that there is no intrinsic block to the formation of long transcripts from the U1 promoter. We infer that the utilization of the histone 3' signals is very efficient, since placing a U1 3' end only 60 nucleotides 3' of the histone 3' end (in the UHHU<sub>156</sub> gene) did not result in accumulation of any transcripts ending at the U1 3' end.

These results do not allow us to determine the mechanism of coupling of the U1 promoter to 3' end formation. There are two likely possibilities: either the promoter determines a specific transcription termination site which is not recognized by RNA polymerase II which has initiated at other promoters; or, alternatively, there is a processing factor which associates with the transcription complex initiating from an snRNA promoter, and processing is tightly coupled with transcription. Clearly the two characterized processing reactions which form the 3' ends of mRNAs, cleavage of the pre-histone

mRNA, or cleavage and polyadenylation of other mRNAs, are posttranscriptional. Thus, if the U1 snRNA 3' ends are formed by processing, it must involve a novel co-transcriptional processing mechanism.

### Acknowledgments

This work was supported by grants GM 27789 and GM 29832 from the NIH to W. F. M.

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

Duane R. Pilch is currently at the National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.

### References

- R. A. Ach and A. M. Weiner (1987), *Mol Cell Biol* 7, 2070-2079.
- C. Birchmeier, D. Schümperli, G. Sconzo, and M. L. Birnstiel (1984), *Proc Natl Acad Sci USA* 81, 1057-1061.
- M. L. Birnstiel, M. Busslinger, and K. Strub (1985), *Cell* 41, 349-359.
- W. G. Chaney, D. R. Howard, J. W. Pollard, S. Sallustio, and P. Stanley (1986), *Somatic Cell Mol Genet* 12, 237-244.
- N. Chodchoy, B. J. Levine, C. Sprecher, A. I. Skoultchi, and W. F. Marzluff (1987), *Mol Cell Biol* 7, 1039-1047.
- N. Chodchoy, N. B. Pandey, and W. F. Marzluff (1991), *Mol Cell Biol* 11, 497-509.
- J. E. Dahlberg and E. Lund (1989), in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (M. L. Birnstiel, ed.), Springer-Verlag, Berlin, pp. 38-70.
- J. E. Dahlberg and E. T. Schenborn (1988), *Nucl Acids Res* 16, 5827-5840.
- O. Gick, A. Krämer, W. Keller, and M. L. Birnstiel (1986), *EMBO J* 5, 1319-1326.
- R. A. Graves, S. E. Wellman, I. M. Chiu, and W. F. Marzluff (1985), *J Mol Biol* 183, 179-194.
- R. A. Graves, N. B. Pandey, N. Chodchoy, and W. F. Marzluff (1987), *Cell* 48, 615-626.
- N. Hernandez and A. M. Weiner (1986), *Cell* 47, 249-258.
- M. M. Hurt, N. Chodchoy, and W. F. Marzluff (1989), *Nucl Acids Res* 17, 8876.
- B. J. Levine, N. Chodchoy, W. F. Marzluff, and A. I. Skoultchi (1987) *Proc Natl Acad Sci USA* 84, 6189-6193.
- B. J. Levine, T. J. Liu, W. F. Marzluff, and A. I. Skoultchi (1988), *Mol Cell Biol* 8, 1887-1895.
- T. J. Liu, B. J. Levine, A. I. Skoultchi, and W. F. Marzluff (1989), *Mol Cell Biol* 9, 3499-3508.

- S. M. Lobo and W. F. Marzluff (1987), *Mol Cell Biol* 7, 4290–4296.
- B. Lüscher and D. Schümperli (1987), *EMBO J* 6, 1721–1726.
- W. F. Marzluff, D. T. Brown, S. Lobo, and S. S. Wang (1983), *Nucl Acids Res* 11, 6255–6270.
- W. F. Marzluff and N. B. Pandey (1988), *TIBS* 13, 49–52.
- N. M. Moussa, A. S. E-L-Din, S. M. Lobo, and W. F. Marzluff (1987), *Nucl Acids Res* 15, 3622.
- K. L. Mowry and J. A. Steitz (1987), *Mol Cell Biol* 7, 1663–1672.
- H. E. Neuman de Vegvar, E. Lund, and J. E. Dahlberg (1986), *Cell* 47, 259–266.
- N. B. Pandey and W. F. Marzluff (1987), *Mol Cell Biol* 7, 4557–4559.
- N. B. Pandey, N. Chodchoy, T. J. Liu, and W. F. Marzluff (1990), *Nucl Acids Res* 18, 3161–3170.
- N. J. Proudfoot (1989), *TIBS* 14, 105–110.
- D. B. Sittman, R. A. Graves, and W. F. Marzluff (1983), *Proc Natl Acad Sci USA* 80, 1849–1853.
- E. Whitelaw, A. Coates, and N. J. Proudfoot (1986), *Nucl Acids Res* 14, 7059–7070.