The EMBO Journal vol.7 no.3 pp.801 - 808, 1988

# Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the *in vitro* 3′ RNA processing reaction

**Matt Cotten, Octavian Gick[1], Alain Vasserot, Gotthold Schaffner and Max L.Birnstiel**

Research Institute of Molecular Pathology, Dr Bohr-Gasse 7, 1030 Vienna, Austria and [1]Institut for Molecular Biology II, University Zurich, Hönggerberg, 8093 Zürich, Switzerland

Communicated by M.L.Birnstiel

**We have made a detailed molecular analysis of the reactions leading to the formation of mature 3′ ends in mammalian histone mRNAs. Using two analytical protocols we have identified an essential sequence motif in the downstream spacer which is consistently present, albeit in diffuse form, in mammalian histone genes. Tampering with this sequence element completely abolishes 3′ processing. However, 3′ cleavage *in vitro*, although at a very much reduced rate, can be detected when the conserved hairpin is deleted from histone precursor mRNAs. U7 snRNA, previously shown to be essential for the maturation of sea urchin histone messages, was isolated from murine cells and the sequence was determined. The ~63-nucleotide, trimethyl-G-capped, murine U7 snRNA possesses a sequence shown in the sea urchin U7 to be required for Sm-precipitability, and like the sea urchin U7, the 3′ end of murine U7 is encased in a hairpin structure. The 5′ sequence of murine U7 exhibits extensive sequence complementarity to the conserved downstream motif of the histone precursor. As expected, oligonucleotide-directed RNase H cleavage of this portion of murine U7 inhibits the *in vitro* processing reaction. These experiments identify a set of specific contacts between mammalian U7 and histone precursor RNA which is indispensable for the maturation reaction.**

*Key words:* RNA maturation/synthetic DNA oligos/cell proliferation/post-transcriptional control

## Introduction

The histone genes coding for the so-called replication variant proteins provide a prime example for the pivotal role that post-transcriptional regulation can play in the eukaryotic cell. While the rate of histone gene transcription fluctuates during the cell cycle, being low in $G_1$ and $G_2$ and high during the S-phase (Heintz *et al.*, 1983), events affecting the stability of histone mRNA in the cytoplasm are at least as important in providing the appropriate levels of translatable histone mRNA (Capasso *et al.*, 1987; Graves *et al.*, 1987). Another, as yet little investigated, aspect of post-transcriptional regulation of histone gene activity is the modulation of the rate of 3′ processing of histone pre-mRNA. It is known, however, that mammalian cells which are arrested in $G_1$ for a protracted period and are thus prevented from cell proliferation exhibit a strong down regulation of 3′ processing (Lüscher and Schümperli, 1987).

The molecular analysis of the 3′ processing reaction in nuclear extracts of HeLa cells has led to the identification of a heat-labile component which cooperates with Sm-subtype snRNP(s) in the post-transcriptional maturation of histone pre-mRNA (Gick *et al.*, 1987). This heat-labile component shows a bimodal molecular distribution on both glycerol gradients and gel filtration columns. It does not contain an epitope for the Sm-antibodies (antibodies from Lupus erythematosus patients) and is fully denatured at 50°C (Gick *et al.*, 1987). In contrast to the persistence of essential snRNP(s), the activity of the heat-labile factor cannot be detected in cells which are arrested and maintained in $G_1$ (Lüscher and Schümperli, 1987).

Here we report on *cis*-acting RNA sequences directing the 3′ processing of the murine H4 pre-mRNA. We identify and characterize the murine U7 snRNA as the nucleic acid moiety of the essential snRNP and suggest how pre-mRNA and U7 RNA sequences interact with one another during 3′ editing of the histone pre-RNA. Although minor snRNAs of mammals have been characterized and sequenced before (reviewed in Reddy and Busch, 1988), U7 snRNA is the only minor snRNA of known function.

## Results

### *About 25 spacer nucleotides beyond the mature 3′ end of the mRNA are required for 3′ processing of the murine H4 pre-mRNA in vitro*

The 3′ terminal segment of the murine histone H4 gene contains all conserved sequences typical of histone genes encoding replication variant histone proteins (Figure 1a). Immediately upstream of the 3′ terminus of the mature RNA species there is an inverted DNA repeat encoding a highly conserved RNA hairpin structure (reviewed in Hentschel and Birnstiel, 1981). Sixteen nucleotides downstream of it there is in this particular gene an AGGAGAGCT sequence, a relatively poor representative of the more common vertebrate spacer motif PuAAAGAGCT (Birnstiel *et al.*, 1985). First, we wished to establish that this downstream sequence element plays an essential role in 3′ processing of this messenger RNA, as it is known to be the case for a similar sequence in sea urchin histone pre-mRNA (Georgiev and Birnstiel, 1985; Schaufele *et al.*, 1986). We used a simple method of generating sequence deletions which obviated much cloning and sequencing work.

Histone H4 transcripts with different spacer extensions were obtained with SP6 *in vitro* transcription of the 3′ terminal sequences of the murine H4 mRNA which had been progressively truncated with *Bal*31 exonuclease. In particular, the SP64-H4-263/57 (Gick *et al.*, 1986) recombinant containing 263 bp and 57 bp upstream and downstream, respectively, of the 3′ cleavage site was used for the resection experiment. The recombinant DNA was cleaved at the *Eco*RI site in the polylinker 3′ to the gene and the DNA was incubated with *Bal*31 exonuclease. The reaction was stopped at

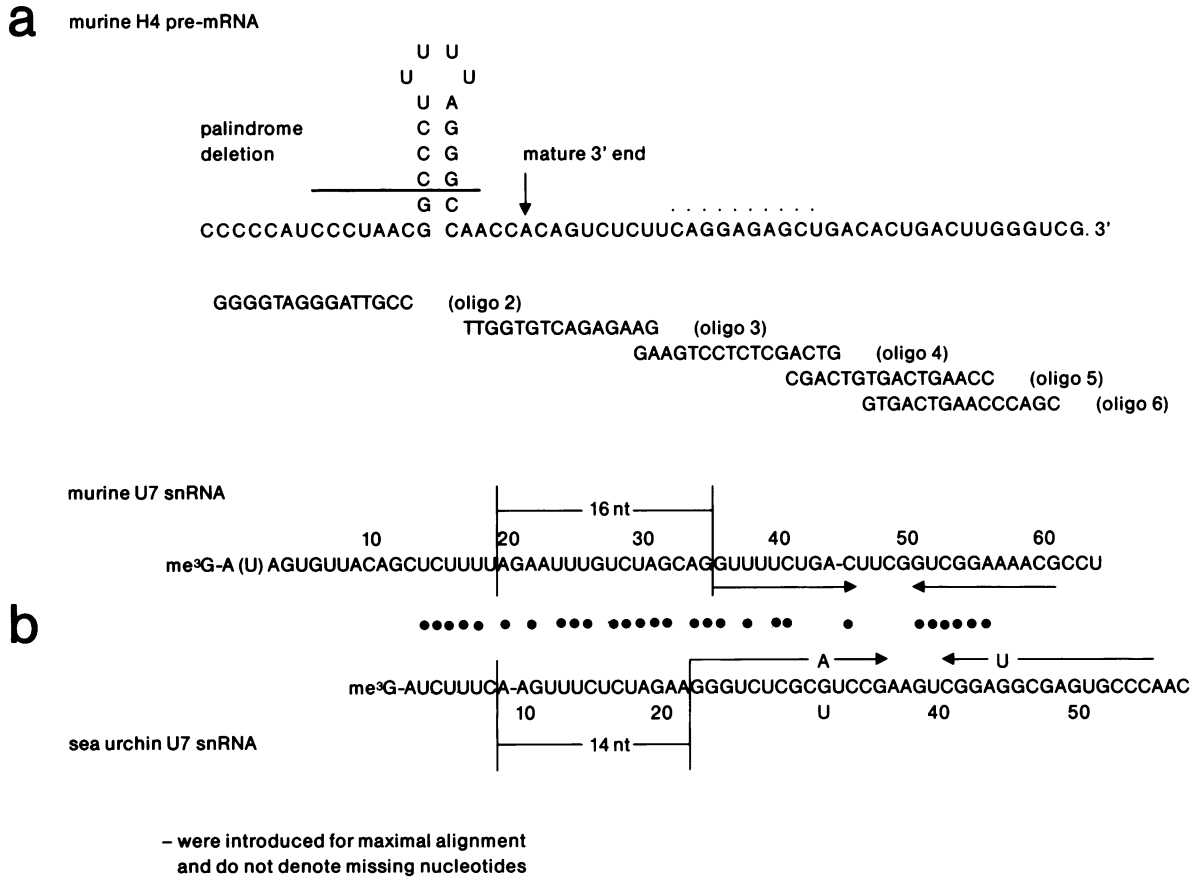


**Fig. 1.** (**a**) Terminal sequences of the murine H4 pre-mRNA flanking the 3′ cleavage site. The palindrome deletion is indicated and complementary oligos used in the experiment in Figure 3 are appropriately aligned. Oligo 1, which is complementary to a sequence 30 nt upstream of the palindrome is not shown. (**b**) Murine and sea urchin U7 snRNAs. Potential palindromes are indicated. The conserved bases are denoted by *. The 16 and 14 nt indicate the number of nucleotides intervening between the pre-mRNA binding sequence and the foot of the palindrome.

different times, the mixtures combined, the DNA isolated and transcribed with SP6 RNA polymerase (Green *et al.*, 1983).

Consecutive RNA size classes were cut out from a preparative PAGE. The RNA recovered from the gel slices was subjected to 3′ processing *in vitro* in a nuclear extract (Gick *et al.*, 1986). The sizes of the precursor RNAs were estimated by taking the mature H4 RNA transcripts on the gels as a fixpoint and using the DNA marker solely as a means to determine the differentials in size. Thus, spacer length near the 3′ end could be determined with an accuracy of a few nucleotides. As can be seen in Figure 2 the 3′ processing reaction is abolished for precursors containing ~10 nucleotides of the spacer beyond the mature 3′ end of the mRNA, i.e. in a precursor in which the downstream conserved sequence has been eliminated (slot 17). When ~20 nucleotides are present, and the conserved motif is included, 3′ maturation ensues, albeit at a somewhat less than maximal rate (slot 16). With RNA molecules containing >30 spacer nucleotides the reaction is maximal (slots 7–15). Very long transcripts are seen to be less efficiently processed but this may be due to nonspecific effects and may result from occlusion phenomena as a consequence of RNA folding.

### *DNA oligonucleotides complementary to the conserved AGGAGAGCU box abolish 3′ processing*

Having determined that the downstream spacer motif is indispensable we wanted to investigate the contribution of the internal RNA sequence. We used a variation of the complementary oligonucleotide approach. Instead of relying on RNase H activity to destroy RNA–DNA complexes, we used the steric hindrance of bound, complementary oligonucleotides to interfere with 3′ processing in the presence of 20 mM EDTA. For this, DNA oligonucleotides 15 or 16 nt long were synthesized which were complementary to preselected regions of the H4 pre-RNA. These were targeted to sequences 30 nt upstream of the palindrome (oligo 1), to sequences immediately 5′ to the terminal hairpin structure (oligo 2), the sequences intervening between the conserved stem and loop structure and the conserved spacer element (oligo 3), the conserved spacer element itself (oligo 4) and two sequences downstream of it (oligos 5 and 6), one with a partial overlap with the conserved spacer sequences (see Figure 1a).

In all cases it was ascertained that the synthetic DNA oligos bound efficiently to the pre-RNA in the processing extract by the simple expedient of adding $Mg^{2+}$ to the processing reaction, a divalent ion which activates the endogenous RNase H. Under these conditions the hybridized RNA is rapidly cleaved by the nuclease to yield the anticipated RNA fragments (see Figure 3, $Mg^{2+}$ slots 1–4). When incubated in 20 mM EDTA, a constituent of the standard processing protocol (Gick *et al.*, 1986), the effect of hybdrid formation between the DNA oligos and the pre-mRNA on 3′ processing

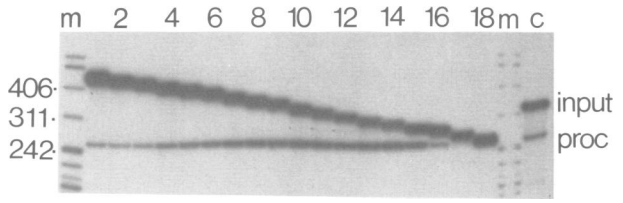


**Fig. 2.** Successive truncation of histone precursor mRNA. pSP64-H4-257/63 (Gick *et al.*, 1986) was linearized at the *Eco*RI site in the polylinker 3′ of the insert, digested with *Bal*31 and used as template for SP6 polymerase. SP6 generated RNAs were separated on a 6% denaturing polyacrylamide gel together with mol. wt markers. Bands corresponding in length to desired 3′ extensions were cut out, extracted and used as precursor for the *in vitro* processing reactions. **Lane C** shows the control processing reaction of the original 257/63 precursor RNA, generated from a *Hin*fI-cut template. The spacer sequences of the precursors used in **lanes 1–18** differ by ~10 nucleotides each, starting with a precursor which just retains the ACCA-sequence (slot 18). Standard processing reactions were performed in HeLa cell nuclear extracts.

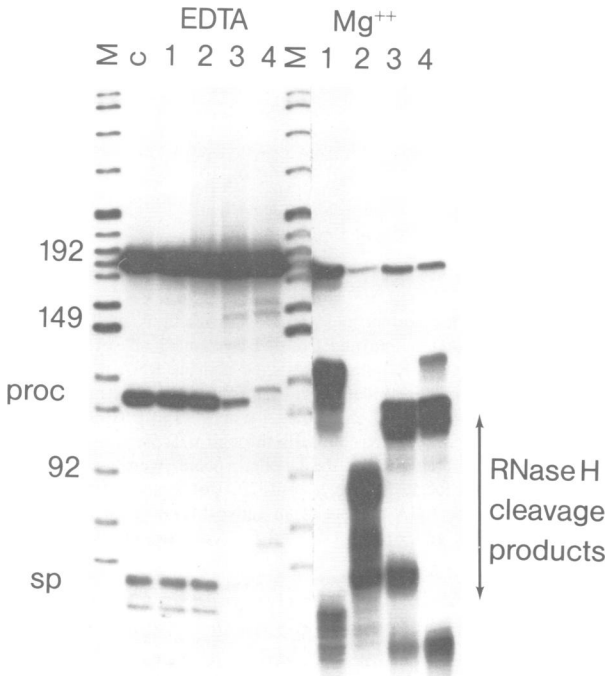


**Fig. 3.** Complementary oligonucleotides are used to probe for regions in the pre-mRNA which are necessary for 3′ processing. Slots 1–4 (EDTA) show standard processing reactions (in HeLa cell nuclear extracts with the 119/70 precursor which had been preincubated for 20 min together with an appropriate 200-fold excess of the oligos depicted in Figure 1 which are numbered in the same way. The slot labelled with 'c' is the minus oligo control. Parallel incubations in the presence of 2 mM $MgCl_2$, instead of 20 mM EDTA (slots 1–4, $Mg^{2+}$), show the expected RNase H cleavage products which confirms that the oligos have indeed hybridized. 'M' denotes *Hpa*II fragments of pBR322 which were included as a size marker.

can be studied.

Oligos 1, 2 and 6 did not diminish the reaction. However, the synthetic DNA which is complementary to the 3′ cleavage site oligo 3 and certainly the oligo 4 complementary to the AGGAGAGCU were inhibitory. In the latter case some minor bands of unusual lengths appear, but their nature was

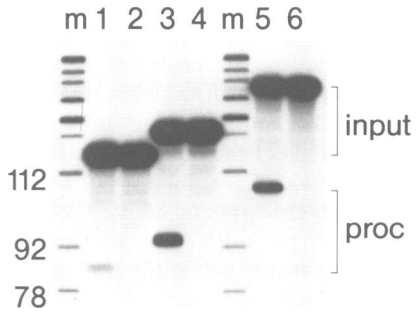


**Fig. 4.** Processing of the palindrome deletion mutants. Standard processing reactions in EBI cell nuclear extract were performed with RNA precursors generated by transcription of H4-119/70 (**slot 6**, input; **slot 5**, processed) H4-87/52 (**slot 4**, input; **slot 3**, processed) and H4-75/52 which is identical to H4-87/52 except for the palindrome deletion (**slot 2**, input; **slot 1**, processed).

not determined. Oligo 5, which overlaps with the last three nucleotides of the spacer core sequence, showed variable results and inhibited the 3′ processing reaction by 20–50% in different experiments (results not shown). These results show that, among those tested, the sequence element centered over the spacer motif is the most important sequence element.

### *3′ cleavage persists even after deletion of the terminal stem-loop structure*

The RNA hairpin in the pre-RNA would not be expected to become available to DNA oligonucleotide hybridization and this called for a different strategy. The contribution of the terminal palindrome to the *in vitro* 3′ processing reaction was determined by deleting it from the pre-RNA (Figure 1a; see Materials and methods).

Precursor RNAs were compared in pairs: first, the wild-type precursors 119/70 and 87/52 (numbers denote nucleotides up- and downstream of the mature 3′ end). 41 nt and 23 nt beyond the spacer motif were included in these two transcripts, respectively. The 119/70 RNA was also compared with the palindrome deletion mutant 75/52 which, like its control (87/52), contains the same number of spacer nucleotides.

Figure 4 shows the relative efficiencies at which 3′ ends are generated for these pre-RNAs. Changing the length of the spacer beyond the conserved motif from 23 to 41 nt has no significant effect on processing efficiency (compare slots 3 and 5). However, the palindrome deletion created a strong down-mutation, by a factor of ~20, but a small amount of cleaved RNA appeared on the gel (Figure 4, slot 1). S1 mapping experiments show that the cleavage site is similar to that of the wild type RNA. However, demonstration that this cleavage produces the genuine 3′ end awaits further experiments.

These initial three sets of experiments demonstrate the overriding importance of the downstream conserved spacer motif for the processing reaction. The next 10 spacer nucleotides are capable of enhancing the reaction (see Figure 2). Deletion of the conserved hairpin creates a strong down-mutation whereas the sequence upstream of the palindrome and the far downstream sequences of the spacer appear to be dispensable for the *in vitro* maturation of the H4 pre-RNA.

### *A 65 nt-long murine RNA is related in sequence to the sea urchin U7 snRNA*

We know from previous experiments that *in vitro* processing of murine H4 pre-RNA is abolished when nuclear processing extracts are depleted of snRNPs with Sm-antibodies (Gick *et al.*, 1986). It has also been demonstrated that a snRNP of the Sm-subtype associates with the 3′ terminal sequences in processing extracts (Mowry and Steitz, 1987). Hence, the 3′ editing of histone pre-mRNAs in the mammalian cell depends on snRNPs, as was previously shown to be the case for the sea urchin system (reviewed in Birnstiel *et al.*, 1985). When processing factors are fractionated according to Krämer (1987) and passed over a DEAE, a heparin Sepharose and Mono Q column, highly purified fractions can be obtained which are capable of supporting 3′ maturation of added histone pre-mRNA (Gick *et al.*, 1987). These active fractions include an ~65 nt long RNA (Krämer, 1987).

An example of the RNA population labelled with [$^{32}$P]-pCp of such an active chromatographic Mono Q fraction is seen in slot 3 of Figure 5. The nominally 65 nt RNA is present as a prominent band together with high mol. wt RNAs, some of which may be contaminating U RNAs and breakdown products thereof. By cutting out the 65 nt RNA band from a preparative gel, the relevant RNA can be purified further (slot 4). The RNA of the heparin Sepharose-bound active fraction, which still contains all snRNPs and is the starting material for the Mono Q fractionation, is shown in slot 2. A comparison of slot 2 and 3 shows the high resolution and separation power afforded by the Mono Q column fractionation. It is also clear that the 65 nt RNA must be present in very low quantities in mouse hybridoma cells.

Because the U7 snRNA of the sea urchin is 57 nt long (Strub *et al.*, 1984) and the nominally 65 nt can be isolated from an active histone pre-mRNA processing fraction, this RNA was the prime candidate for representing the murine homologue of the sea urchin U7 snRNA. The possible identity of this RNA with its sea urchin counterpart was further strengthened when we discovered that this RNA could be quantitatively precipitated (results not shown) with anti-me$_3$G-cap antibodies (a kind gift of R.Lührmann).

The putative U7 snRNA, isolated as shown in Figure 5, slot 4, in [$^{32}$P]pCp labelled form was subjected to both chemical (Peattie, 1979) and enzymatic RNA sequencing. The enzymatic sequence ladder, when compared with a nucleotide ladder generated by treatment of the pCp-labelled RNA with KOH, allowed exact positioning of the nucleotides relative to the 3′ terminus. The 5′ portion of the RNA sequences were confirmed by retrotranscription of the RNA from DNA primers, using the dideoxy method. The 5′ terminal A(U)A, shown in Figure 1b, could not be determined with confidence in our sequencing experiments. Sequencing of the U7 snRNA genes will be required to obtain a definitive sequence. A similar, but not identical sequence has also been obtained by D.Soldati and D.Schümperli (personal communication); this could mean that murine U7 snRNA is polymorphic as is suggested to be the case with the sea urchin homologue (De Lorenzi *et al.*, 1986).

The sequences of the murine and of the sea urchin U7 snRNA (Strub *et al.*, 1984) are shown and compared in Figure 1b. Murine and sea urchin U7 RNAs show a considerable degree of sequence homology, especially in the 5′

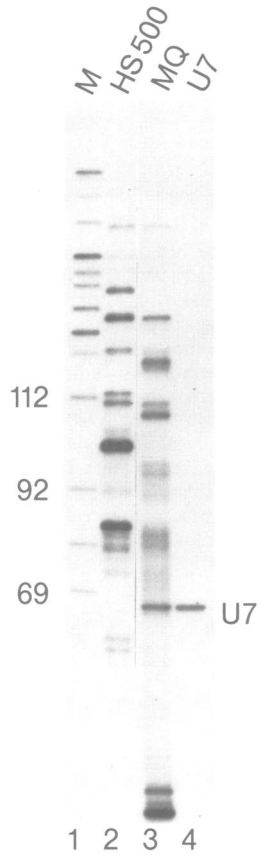


**Fig. 5.** RNA species present in 3′ processing fractions. HS500 and an active Mono Q fraction, purified from EBI cell nuclear extract, were treated with proteinase K, phenol extracted and the RNA present was recovered by ethanol precipitation in the presence of 0.02 mg/ml glycogen. The RNA was 3′ end-labelled with [$^{32}$P]pCp (England *et al.*, 1980), resolved by electrophoresis through a 12% acrylamide/8.3 M urea gel and the labelled species were visualized by autoradiography. **Slot 1**, *Hpa*II-cut pBR322; **slot 2**, the RNA in HS500; **lane 3**, the RNA present in an active Mono Q fraction; **slot 4**, a sample of the gel-purified U7 used for sequence analysis.

half of the molecule. The 3′ portion of both RNA molecules can be folded into a palindrome. In the case of murine U7 snRNA there are no mismatched base pairs and only one GU base pair in the hairpin. It differs in sequence from the comparable stem-loop in the sea urchin homologue. The murine U7 snRNA, depending on whether the U of the initial A(U)A is included or not, is 62 or 63 nt long respectively. The sequences of primary interest, for the present, are located at the 5′ ends of both snRNAs. They contain prominent base complementarities to the spacer motif of their respective mRNA species (see Figure 7a).

### *The identity of the murine U7 snRNA as an essential component of the 3′ processing machinery is confirmed by DNA oligonucleotide-mediated truncation of the U7 snRNA*

If the sequenced RNA is truly involved in histone messenger processing, then one should be able to demonstrate an inhibition of *in vitro* 3′ maturation of histone pre-mRNA when the 5′ terminal sequences are challenged with a DNA oligo complementary to nt 4–21, followed by RNase H digestion. A complete inhibition of processing activity is indeed

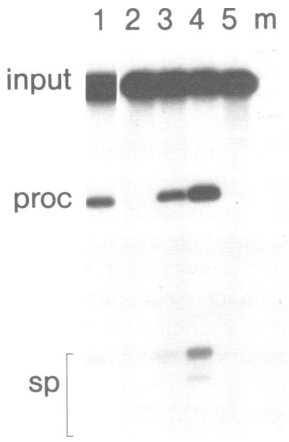


**Fig. 6.** Inhibition of 3′ processing with oligonucleotides complementary to murine U7. EBI cell nuclear extracts were treated with RNase H (0.8 units/0.015 ml extract) in the presence of 6 mM magnesium and various oligonucleotides for 2 h at 30°C. EDTA was added (to 20 mM) and a standard processing reaction was performed. **Slot 5**, processing substrate, no nuclear extract; **slot 4**, standard processing reaction; **slot 3**, processing reaction with an extract pretreated with RNase H and magnesium, but no oligonucleotide; **slot 2**, processing reaction with an extract treated with RNase H in the presence of an oligonucleotide complementary to nucleotides 4–21 of murine U7 snRNA; **slot 1**, processing reaction with an extract treated with RNase H in the presence of an oligonucleotide complementary to nucleotides 27–35 of murine U7 snRNA.

observed with such a synthetic DNA molecule (Figure 6, compare slot 2 and control in slot 4). Incubation with RNase H alone in the absence of the complementary oligo (slot 3), or in the presence of an oligo targeted to an internal sequence of the U7 snRNA (nt 27–35, slot 1) have only minor effects on the processing activity. These experiments provide further proof that with isolation of the ~63 nt-long RNA from the mouse we have identified the murine U7 snRNA.

## Discussion

### *The sequences in control of 3′ editing of histone pre-mRNA show functional equivalence in both mouse and sea urchin*

The results presented here confirm essential points of our previous studies on sea urchin histone 3′ processing analysed, in the main, by frog oocyte injection experiments (reviewed by Birnstiel and Schaufele, 1988). The importance of the downstream conserved spacer sequences is demonstrated again (cf. Georgiev and Birnstiel, 1985). The sequences immediately upstream of the palindrome are seen to be dispensable (cf. Birchmeier *et al.*, 1982). This contrasts with results of Stauber *et al.* (1986). The sequences between the conserved histone RNA hairpin and the downstream spacer element are found to contribute to the efficiency of the 3′ processing reaction (cf. Georgiev and Birnstiel, 1985). The spacer immediately beyond the downstream CAAGAAAGA of the sea urchin (Georgiev and Birnstiel, 1985) and the AGGAGAGCU of the mouse H4 (this paper) can increase the efficiency of 3′ processing, but this effect seems less pronounced *in vitro* (this paper) than *in vivo* (Birchmeier *et al.*, 1983; cf. Stauber *et al.*, 1986).

The persistence of some 3′ cleavage at the correct site after complexing of the cleavage site with a complementary DNA oligonucleotide might mean some cleavage is possible even in an RNA/DNA hybrid. It must also be considered that, in contrast to the standard DNA oligo targeted destruction of the RNA sequences, the simple hybridization of short DNA sequences may not provide the same all or none effect. Possibly, the hybrids that form are subject to some dissociation or 'breathing' in the course of the processing reaction. If this were true, then the cleavage site could become transiently available following dissociation of the hybrid. There is persuasive evidence, discussed below, that the terminal GCU of the spacer motif is an indispensable part (especially for the murine H4 pre-mRNA) of the U7 snRNA/histone pre-mRNA contacts. The above considered dynamic aspects of the RNA/DNA hybrids might provide a ready explanation of why a synthetic DNA overlapping with the terminal GCU does not effectively block the processing reaction. In this particular case, the DNA oligo might also be in direct competition with the U7 snRNA binding to the histone pre-mRNA.

There is compelling evidence that the terminal stem-loop structure contributes to the regulation of histone mRNA stability (Georgiev *et al.*, 1984; Capasso *et al.*, 1987; Graves *et al.*, 1987) and to histone pre-mRNA processing (Birchmeier *et al.*, 1984; Georgiev and Birnstiel, 1985; Stauber *et al.*, 1986). Here we show that deletion of the palindrome still allows the generation of 3′ ends *in vitro*, albeit at a much reduced rate. Previously, the palindrome was held to be an absolute requirement for 3′ RNA processing (reviewed by Birnstiel *et al.*, 1985). It is possible that a reduced stability *in vivo* of the processed mRNA lacking the palindrome would account for the failure to observe the small amount of product which is generated even from this mutant. Conversely, deletion of the AGGAGAGCU motif quite clearly abolishes all 3′ maturation both *in vivo* and *in vitro*. Thus, while both elements are important for directing the 3′ maturation reaction, a pre-eminence of the spacer motif is clear, at least for the *in vitro* reaction.

### *Some U7 snRNA sequences are conserved between sea urchins and mice*

In keeping with their involvement in the 3′ editing of histone pre-mRNAs one finds that the 5′ terminal sequences of the U7 snRNAs are related to one another and both show a potential for basepairing to the conserved spacer motif on their respective histone mRNAs. In the sea urchin this 5′ sequence is readily digestible with micrococcal nuclease and presumably available for pre-mRNA binding (Gilmartin *et al.*, 1988). The vertebrate RNA is strikingly different, however, in that it has an added 11–12 nt at the 5′ end in keeping with the proposal of Turner *et al.* (1983) that contacts between the vertebrate U7 snRNA and the histone pre-mRNA should include nucleotides downstream of the PuAAAGA spacer motif.

It was suggested from the mutagenesis of sea urchin U7 snRNA that the nucleotide stretch AGUUUCUCUAGA (nt 8–21) includes an Sm-binding site (Gilmartin *et al.*, 1988). Interestingly, the topologically related sequence (nt 19–32) is conserved in the murine U7 snRNA. The latter, but not the former, RNA includes a GAAUUUGU which is a vestige of the vertebrate Sm-binding site which usually takes the form of PuA$(U)_n$ GPu (Mattaj and De Robertis, 1985) where $n$ is rarely 3, but mostly 4–6. As shown in the experiment of Figure 6, slot 1, this sequence, although present

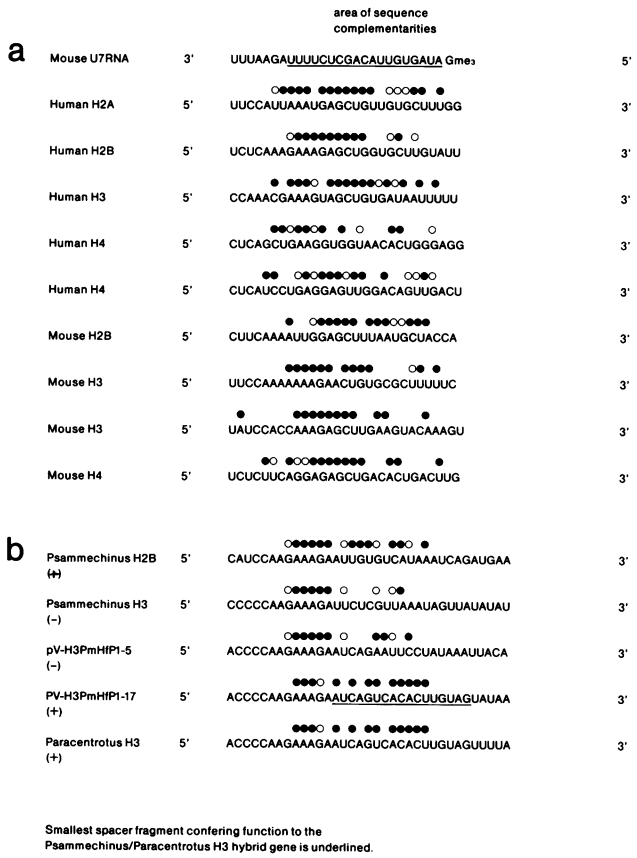


**Fig. 7.** Base complementarities between murine U7 snRNA and the downstream spacer sequences of (**a**) mammalian and (**b**) sea urchin histone genes. The sequences were taken from Wells (1986) and Schaufele *et al.* (1987). To allow for a maximal fit the *Paracentrotus* H3 sequences were displaced by one nucleotide. ● AU, GC; ○ GU pairs. A (+) indicates that the transcript is processed in frog oocytes without the addition of sea urchin U7 snRNA, a (−) indicates an absence of processing without the addition of sea urchin U7 snRNA.

in linear form, is bound up within the snRNP structure and is not open to a concerted attack by complementary DNA nucleotides and RNase H. The same RNA segment is protected from micrococcal nuclease digestion in the sea urchin snRNP (Gilmartin *et al.*, 1988). It seems likely therefore that the same or similar, rather unique, sequence in both sea urchin and mouse is bound up with proteins and mediates the Sm-precipitability of the snRNP particles. Such a function would also adequately explain why this region has no sequence complementarities to histone pre-mRNA.

In contrast to the sea urchin homologue the potential palindrome of the murine U7 snRNA has no mismatched base pairs, although it contains one G.U pair. It differs drastically in sequence from its sea urchin counterpart (Strub *et al.*, 1984). It has previously been shown that, in the sea urchin, such a terminal palindrome is essential for formation of the Sm-precipitable RNA particle and for 3′ processing. However, its sequence could be changed at will with impunity as long as its double-stranded nature was preserved (Gilmartin *et al.*, 1988). It is therefore perhaps not surprising that strong evolutionary constraints on the primary sequence of the palindrome are not in evidence from inspection of the murine and sea urchin U7 snRNA sequence.

### *The sequences of the murine U7 snRNA reveal hyphenated sequence complementarities to the conserved spacer motif of man/mouse (and sea urchins)*

The 5′ terminal sequences of the murine U7 snRNA show rather diffuse base complementarities to all known mammalian histone mRNA which are specifically expressed in proliferating cells. A compilation of the potential basepairings is given in Figure 7a, where human genes are also included. Their addition seems warranted because the human U7 snRNA possesses an identical binding sequence (K.Mowry and J.Steitz, personal communication). These base complementarities concern, in the main, nucleotides 9−18 of the U7 snRNA. In some instances the complementarities extend further upstream to the 5′ end of the U7 snRNA, but downstream sequences appear not to be significantly involved. Therefore, the U7 snRNA histone pre-mRNA complementarities of mammals map primarily over the universal spacer core sequence PuAAAGAGCU (Birnstiel *et al.*, 1985).

Although in all cases shown in Figure 7a the universal core is part of the binding sequence, even here there is some diversity as to the juxtaposed complementary bases. Since the hybridizing sequence is both conserved in mouse (this paper) and man (K.Mowry and J.Steitz, personal communication) it would appear likely that a single U7 snRNA sequence is capable of interacting with different histone pre-mRNAs. Therefore, the hybrids that form can, in general, be rather degenerate. However, in all cases, there is a reasonable stretch of uninterrupted complementarities which could contribute an average of −16 kcal to the stabilization of a hybrid structure. That RNA/RNA hybridization is indeed a significant part of the 3′ processing reaction is strongly suggested from the study of compensatory mutants in the sea urchin system (Schaufele *et al.*, 1986) where a nearly invariant short GAAAGA sequence with six continuous non-degenerate base pairs provide sufficient specificity (probably in conjunction with other contacts yet to be specified) for 3′ processing.

If it is assumed that an identical, or at least similar, binding sequence exists in *Xenopus* U7 snRNA as in mammals, the curious inability for processing of the H3 pre-mRNA of *Psammechinus* H3 pre-mRNA, but not of the related *Paracentrotus* H3 pre-mRNA, in *Xenopus* oocytes can be readily explained at a molecular level. The distinctive behaviour of these two RNAs was previously traced to a 15 nt-long RNA segment immediately following the (sea urchin type) conserved CAAGAAAGA spacer motif (Schaufele and Birnstiel, 1987). The sequence disposition in the two pre-RNAs and their mutants used to delimit the RNA segment which imparts RNA processing activity is shown in Figure 7b.

It is clear that the sea urchin histone precursor RNAs do not have the vertebrate GCU and yet most of them are processed in the frog oocytes. In the *Psammechinus* H2B pre-mRNA and the *Paracentrotus* H3 pre-mRNA this appears compensated by additional (presumptive) contact sites for the vertebrate U7 snRNA. However, the processing deficient *Psammechinus* H3 pre-mRNA has a great deficit of bases capable of forming base pairs with such a U7 snRNA. The quite varied disposition of sequence complementarities between, for instance, the processing active *Psammechinus*

H2B and the *Paracentrotus* H3 gene provide further evidence that there need be no uniquely defined binding sequence, rather that many different pairing constellations may be acceptable, as long as they contain a sufficient number of contact sites. Thus a situation similar to the recognition of the 5′ splice site by the U1 snRNA during RNA splicing prevails (Zhuang and Weinger, 1986), where the area as a whole rather than an exact sequence is recognized (Aebi *et al.*, 1987).

Some sequence complementarities between the U7 snRNA and the RNA stem-loop of the histone mRNA are present as well. However, there is now compelling evidence for the sea urchin that such sequence complementarities are not an essential part of the processing reaction (Schaufele and Vasserot, 1988).

### *The increased distance between the histone palindrome and the downstream spacer motif may have been compensated in evolution by an alteration in the U7 snRNA sequence*

It was noted before that the distance between the base of the histone palindrome and the downstream spacer motif may differ between species but it is rigidly conserved within the species (Birnstiel *et al.*, 1985). Thus, there are 13 or 14 nt in sea urchin, and 15 or 16 nt in mouse and man. This evolutionary rigidity may arise from a need to keep both palindrome and downstream sequence in a defined topological relationship to one another. This interpretation is supported by experiments in which alterations of the number of nucleotides intervening between these *cis*-acting elements creates strong down-mutations (Georgiev and Birnstiel, 1985). Even addition of only two intervening nucleotides in the sea urchin pre-mRNA was found to reduce the processing rate by as much as 50% in frog oocyte injection experiments (Schaufele, 1987).

While some flexibility, within relatively narrow limits, for the arrangement of these two key processing elements exists, it is nevertheless pleasing to note that the nucleotides between the U7 snRNA palindromes and the sequences binding to the histone spacer core sequence number 14 in the sea urchin and 16 in mammals. Therefore, when the U7 snRNAs in both species are aligned relative to histone pre-mRNA in a processing configuration, the U7 snRNA palindromes in both species are brought into exact juxtaposition with their respective histone palindromes, making interactions between the structures possible, at least in theory.

From a large body of mutagenesis experiments it is clear that such interactions, if they exist, are not likely to depend on Watson–Crick base pairing. It is a curious feature for both U7 snRNA and histone pre-mRNA palindromes that while their hairpin structure must be maintained for efficient 3′ processing to occur, their primary sequence can be changed at will without abolishing 3′ cleavage of the pre-mRNA (Gilmartin *et al.*, 1988; Schaufele and Vasserot, 1988).

The structural data presented here say nothing about the role of the heat-labile factor during RNA 3′ processing. Armed with the above information we are now in a position to determine how this factor fits into the general scheme which regulates the production of mature 3′ ends from histone pre-mRNAs.

## Materials and methods

Preparation of nuclear extracts from HeLa cells or EBI cells (a mouse hybridoma line), SP6 transcriptions of pSP65-H4-119/70 and *in vitro* pre-mRNA processing reactions were as described earlier (Gick *et al.*, 1986) with the exception that ATP and creatine phosphate were omitted in the processing mixture. *Escherichia coli* tRNA was used as a carrier at 0.15 mg/ml. Chromatographic purification of processing activity was performed as described by Krämer (1987).

#### *RNA sequence analysis*

RNA was obtained from Mono Q-purified processing extracts (derived from EBI cells) by proteinase K digestion and phenol/chloroform extraction followed by ethanol precipitation in the presence of 20 μg/ml glycogen. Labelling of the RNAs with [$^{32}$P]pCp was essentially as described by England *et al.* (1980). RNA to be sequenced was resolved by electrophoresis through a 12% acrylamide, 8.3 M urea gel, visualized by a brief exposure to X-ray film, the appropriate band was excised and the RNA in the gel slice was eluted into 0.4 ml HEB [0.75 M ammonium acetate, 10 mM magnesium acetate, 1% (v/v) phenol, 0.1% (w/v) SDS, 0.1 mM EDTA] with an overnight incubation in an Eppendorf vibrator. The RNA was again extracted with phenol/chloroform and precipitated with ethanol in the presence of 0.005 mg tRNA. Chemical sequencing of the RNA was performed as described by Peattie (1979). The 3′ end sequence of the U7 snRNA was determined by digesting 3′ end-labelled RNA samples nearly to completion with either RNase T1, RNase U2 (8 or 4 units, in 20 mM Na-citrate, pH 3.5, 1 mM EDTA, 8 M urea), RNase CL3 (0.08 units, in 20 mM Tris, pH 7.5, 8 M urea) or micrococcal nuclease (10 units, in 20 mM Tris, pH 7.5, 1 mM $CaCl_2$, 8 M urea) at 50°C for 90 min in the presence of 0.005 mg unlabelled carrier tRNA. The enzymatic digestions were performed in a 0.003 ml volume. Digested samples were resolved on a 20% acrylamide gel with a sample of 3′ end-labelled U7 partially hydrolysed with KOH (150 mM, 12 min, 42°C) as a marker.

#### *Plasmid constructions*

Plasmids −52/+28 (Stauber *et al.*, 1986) and H4-119/70 (Gick *et al.*, 1986) were cleaved with *Hae*III. Two *Hae*III sites are contained in the stem of the conserved terminal palindrome. Cleavage with *Hae*III thus cuts out 12 of the 16 palindrome nucleotides. The 5′ portion of the palindrome minus construct H4-75/52 is obtained by cutting the *Hae*III digested −52/+28 with *Eco*RI and isolating the 70 bp fragment spanning the sequence directly upstream of the palindrome. The 3′ portion is contributed by the 73 bp fragment resulting from *Bam*HI cleavage of the *Hae*III digested H4-119/70. These pieces were ligated into an *Eco*RI, *Bam*HI cut pSP 65 to yield the palindrome minus plasmid. To generate the palindrome plus control H4-87/52, the appropriate *Eco*RI or *Bam*HI fragments from *Alu*I digested −52/+28 and H4-119/70 were obtained and ligated into *Eco*RI, *Bam*HI cut pSP 65.

#### *Oligonucleotide inhibition experiments*

Oligonucleotides were purified on 12% acrylamide, 8.3 M urea, 1 × TBE gels, excised, extracted, phenolized, precipitated and passed over a G-25 column (equilibrated with $H_2O$) to remove urea and salt. The lyophilized oligonucleotides were dissolved in TE at a concentration of 50 nmol/ml and added at a 200-fold molar excess over precursor RNA in the processing reactions. To control for proper hybridization parallel reactions were performed in the presence of 2 mM $MgCl_2$ instead of 20 mM EDTA and for 20 min only. Endogenous RNase H present in the extracts destroys the hybridized portions of the RNA precursor, but leaves the unpaired 5′ and 3′ portions intact.

## Acknowledgements

We are grateful to Eddie De Robertis (L.A.) and Reinhard Lührmann (Berlin) for their gifts of antisera. We thank Jennifer Mitcham for her assistance with retrotranscription. We also thank Joan Steitz and Daniel Schümperli for sharing their U7 sequence information prior to publication. We are grateful to Marianne Vertes for her assistance in typing the manuscript.

## References

Aebi,M., Hornig,H. and Weissmann,C. (1987) *Cell*, **50**, 237–246.
Birchmeier,C., Grosschedl,R. and Birnstiel,M.L. (1982) *Cell*, **28**, 739–745.
Birchmeier,C., Folk,W. and Birnstiel,M.L. (1983) *Cell*, **35**, 433–440.

Birchmeier,C., Schümperli,D., Sconzo,G. and Birnstiel,M.L. (1984) *Proc. Natl. Acad. Sci. USA,* **81**, 1057–1061.
Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell,* **41**, 349–359.
Birnstiel,M.L. and Schaufele,F. (1988) In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor snRNAs.* Springer Verlag, Berlin, pp. 155–182.
Busslinger,M., Portmann,R. and Birnstiel,M.L. (1979) *Nucleic Acids Res.,* **6**, 2997–3008.
Capasso,O., Bleecker,G.C. and Heintz,N. (1987) *EMBO J.,* **6**, 1825–1832.
De Lorenzi,M., Rohrer,U. and Birnstiel,M.L. (1986) *Proc. Natl. Acad. Sci. USA,* **83**, 3243–3247.
England,T.E., Bruce,A.G. and Uhlenbeck,O.C. (1980) *Methods Enzymol.,* **65**, 65–74.
Georgiev,O. and Birnstiel,M.L. (1985) *EMBO J.,* **4**, 481–489.
Georgiev,O., Mous,J. and Birnstiel,M.L. (1984) *Nucleic Acids Res.,* **12**, 8539–8551.
Gick,O., Krämer,A., Keller,W. and Birnstiel,M.L. (1986) *EMBO J.,* **5**, 1319–1326.
Gick,O., ,Krämer,A., Vasserot,A. and Birnstiel,M.L. (1987) *Proc. Natl. Acad. Sci. USA,* **84**, 8937–8940.
Gilmartin,M.G., Schaufele,F., Schaffner,G. and Birnstiel,M.L. (1988) *Mol. Cell. Biol.,* in press.
Graves,R.A., Pandey,N.B., Chodchoy,N. and Marzluff,W.F. (1987) *Cell,* **48**, 615–626.
Green,M.R., Maniatis,T. and Melton,D.A. (1983) *Cell,* **32**, 681–694.
Heintz,N., Sive,H.L. and Roeder,R.G. (1983) *Mol. Cell. Biol.,* **3**, 539–550.
Hentschel,C. and Birnstiel,M. (1981) *Cell,* **25**, 301–313.
Krämer,A.R. (1987) *Proc. Natl. Acad. Sci. USA,* **84**, 4159–4168.
Krämer,A.R., Keller,W., Appel,B. and Lührmann,R. (1984) *Cell,* **38**, 299–307.
Lüscher,B. and Schümperli,D. (1987) *EMBO J.,* **6**, 1721–1727.
Lüscher,B., Stauber,C., Schindler,R. and Schümperli,D. (1985) *Proc. Natl. Acad. Sci. USA,* **82**, 4389–4393.
Mattaj,I.W. and De Robertis,E.M. (1985) *Cell,* **46**, 111–118.
Mowry,K.L. and Steitz,J.A. (1987) *Mol. Cell. Biol.,* **7**, 1663–1672.
Peattie,D.A. (1979) *Proc. Natl. Acad. Sci. USA,* **76**, 1760–1764.
Reddy,R. and Busch,H. (1988) In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor snRNAs.* Springer Verlag, Berlin, pp. 1–37.
Schaufele,F. (1987) Ph.D. Thesis, University of Zürich.
Schaufele,F. and Birnstiel,M.L. (1987) *Nucleic Acids Res.,* **15**, 8305–8317.
Schaufele,F. and Vasserot,A. (1988) *Proc. Natl. Acad. Sci. USA,* **85** (in press).
Schaufele,F., Gilmartin,G.M., Bannwarth,W. and Birnstiel,M.L. (1986) *Nature,* **323**, 777–781.
Stauber,C., Lüscher,B., Eckner,R., Lötscher,E. and Schümperli,D. (1986) *EMBO J.,* **5**, 3297–3303.
Strub,K. and Birnstiel,M.L. (1986) *EMBO J.,* **5**, 1675–1682.
Strub,K., Galli,G., Busslinger,M. and Birnstiel,M.L. (1984) *EMBO J.,* **3**, 2801–2807.
Turner,P.C., Aldridge,T.C., Woodland,H.R. and Old,R.W. (1983) *Nucleic Acids Res.,* **11**, 4093–4107.
Wells,D.E. (1986) *Nucleic Acids Res.,* **14**, r119–r149.
Zhuang,Y. and Weiner,A.M. (1986) *Cell,* **46**, 827–835.

*Received on December 11, 1987; accepted on December 31, 1987*