Volume 14 Number 17 1986 Nucleic Acids Research

Globin gene transcripts can utilize histone gene 3' end processing signals

E.Whitelaw, A.Coates and N.J.Proudfoot

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Received 24 June 1986; Accepted 6 August 1986

ABSTRACT

Deletion of the poly(A) site from the human α globin gene results in a defective gene that produces very little stable mRNA as compared to the intact gene, presumably due to the instability of the mRNA. However, if the α poly(A) site is replaced by mouse histone H4 3' end processing signals, significant levels of hybrid α/H4 mRNA are obtained and the transcripts formed are cytoplasmic and poly(A)$^-$. When both mouse histone 3' end processing signals and the α globin poly(A) site signals are placed in tandem after the α globin gene promoter and coding sequence, the α poly(A) site signals are utilized exclusively. These results show that the histone 3' end processing signals can function independently of the histone promoter and the transcripts which are normally polyadenylated (α globin) can be stabilized by a poly(A)$^-$ histone mRNA 3' terminus. Furthermore, these results show that the histone 3' end processing signals are less efficient than the α globin poly(A) site signals, if the two are placed in direct competition.

INTRODUCTION

The differential expression of eukaryotic genes is an essential feature of development. Much of this gene regulation is likely to be associated with promoter selection (1, for review). However, another important mechanism open to eukaryotes for regulating gene expression may be the degree of stability or half life of a gene transcript. In the case of genes transcribed by RNA polymerase II (polII genes) the formation of the mRNA's 3' terminus is often associated with mRNA stabilization (2, for review). In particular the addition of a 3' terminal poly(A) tail has been shown to confer stability on mRNA in a number of different experimental systems (3). Two unusual types of polII gene transcripts have been shown to lack poly(A) tails; small nuclear RNAs (4) and histone mRNAs (5). Both of these mRNAs have 3' terminal hairpin loops which may confer some degree of stabilisation on the associated RNA molecule.

We wished to investigate the role of mRNA 3' end formation in RNA stability by testing whether or not a normally polyadenylated mRNA (human α globin) can utilise a histone mRNA 3' end processing signal in place of its own poly(A) site. We demonstrate in these studies that the mouse histone H4 gene 3' processing signals can

© IRL Press Limited, Oxford, England.

function when exchanged for the α globin poly(A) site even though reduced levels of mRNA are obtained. The resulting globin-histone hybrid mRNA is poly(A)$^-$. Interestingly when the normal α globin poly(A) site is positioned downstream from the histone 3' processing signals, no histone 3' ends are detectable but rather normal levels of α globin transcripts are obtained that read through the H4 insert to form polyadenylated 3' termini.

METHODS

1. Gene Constructs

α pSVod. The plasmid α pSVod contains the 1.6 Kb PstI fragment containing the human α1 globin gene (see Fig.1). This fragment has been inserted between the PvuI and PstI sites of the expression vector pSVod. pSVod contains the pBR322 replication origin, the tetracycline gene and the SV40 replication origin (6).

H4 pSVed. The plasmid H4 pSVed contains the 690 bp EcoR1 to HinfI fragment containing the entire mouse H4 histone gene (see Fig.1). This fragment has a BamHI linker and a HindIII linker at its 5' and 3' end respectively (7). This fragment was cloned into the expression vector pSVed between the BamHI and HindIII sites. pSVed contains the pBR322 replication origin and the SV40 replication origin and enhancer (8).

H4* pSVed. This plasmid is derived from H4 pSVed. A BglII linker was inserted into the XmnI site within the H4 histone coding sequence (see Fig.2).

αΔBP pSVod. This plasmid is derived from α pSVod. The 300 bp BstEII to PvuII fragment containing the α poly(A) site has been deleted (see Fig.1).

α H4 pSVod. This plasmid is derived from α pSVod. The 300 bp BstEII to PvuII fragment containing the α poly(A) site has been removed and has been replaced by the 250 bp XmnI to HindIII fragment containing the 3' end of the mouse H4 histone gene (see Fig.1).

α/H4/α pSVod. This plasmid is derived from α pSVod. The 250 bp XmnI to HindIII fragment from the 3' end of the mouse H4 histone gene has been inserted into the α globin gene at the BstEII site.

pRBSV. This plasmid was a gift from Dr F. Grosveld. The rabbit β globin gene has been inserted into a plasmid containing the SV40 Ori and enhancer sequences and the large T antigen sequences (9). The cotransfection of this plasmid allows for the replication in HeLa cells of any plasmid containing the SV40 origin.

2. Transient Expression

Transfections into HeLa cells were carried out as previously described (6). Plasmid DNA was calcium phosphate precipitated and added to subconfluent dishes of HeLa cells. After 16 hours, the medium was changed and the cells were allowed to

grow for another 36 hours. All transfections were carried out in the presence of the plasmid, pRBSV, containing the rabbit β globin gene, the SV40 origin and enhancer and the large T antigen sequences (9).

3. RNA Purification

The cells were harvested and the RNA was purified as previously described (6). Essentially, cells were lysed in NP40 detergent buffer and the cytoplasmic and nuclear fractions were separated by centrifugation through a sucrose cushion. Following incubation with proteinase K, cytoplasmic RNA was purified by phenol/chloroform extraction and ethanol precipitation. Nuclear RNA was prepared by including a DNase digestion after the proteinase K step. Poly(A) selection was carried out by standard procedures using an oligo(dT) cellulose column (10).

RNA Mapping

S1 nuclease. Probe DNAs (either double or single strand) were annealed to cytoplasmic RNAs (10-20 mg) in 30 µl of 80% formamide, 0.04 M PIPES pH 6.4, 0.4 M NaCl, 0.1 mM EDTA by denaturing at 80°C for 10 min and then hybridizing at 53°C (double stranded probe) or 30°C (single stranded probe) overnight. 300 µl of ice-cold S1 buffer (0.25 M NaCl, 0.03 M NaOAc pH 4.6, 2 mM $ZnSO_4$, 50 µg/ml denatured sonicated carrier DNA plus S1 nuclease (3000 units) was added and incubated for 1 hour at 30°C. S1 reactions were ethanol precipitated and fractionated on denaturing 7 M urea polyacrylamide gels.

Primer Extension. Labelled DNA primer and cold RNA were annealed in 10 µl of 10 mM PIPES pH 6.4, 0.4 M NaCl at 80°C for 10 minutes and then hybridized at 63°C overnight. 50 µl of reverse transcriptase buffer [50 mM Tris, pH 8.2, 10 mM DTT, 6 mM $MgCl_2$, 0.5 mM dATP, dCTP, dTTP, dGTP plus reverse transcriptase (5 units)] were added to the hybridizations and incubated at 42°C for 1 hour. The reactions were ethanol precipitated, RNase-treated and then fractionated by electrophoresis on 7 M urea polyacrylamide gels.

Exonuclease VII digestion. Hybridizations were carried out as for the S1 experiments. 0.5 ml of exonuclease VII buffer (30 mM KCl, 10 mM Tris pH 7.8 and 10 mM EDTA) with 4 µl exonuclease VII (BRL) was added to the hybridizations and incubated at 37°C for 2 hours. The reactions were then ethanol precipitated and fractionated on denaturing polyacrylamide gels.

RESULTS

Mouse histone H4 3' end processing in a HeLa cell transient expression system

To begin these studies we wished to ascertain whether or not the mouse histone H4 gene (7) could form mRNA with authentic 3' ends when transfected into human HeLa cells. This was an essential point to establish since it has been shown that the

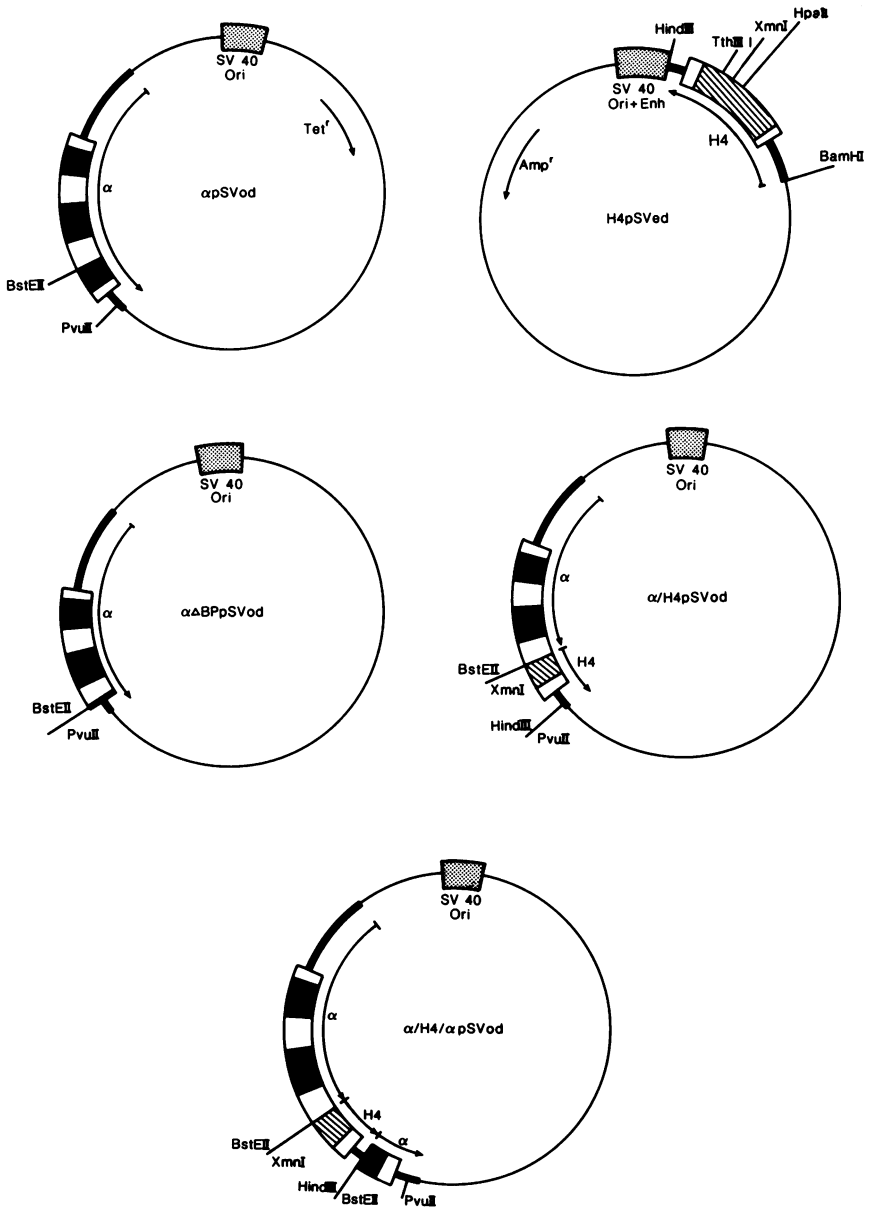

Figure 1. Maps of the plasmids α pSVod, H4 pSVed, αΔBP pSVod, α/H4 pSVod, α/H4/α pSVod. Thin lines indicate pBR322 sequence, thick lines indicate 5' and 3' flanking sequence, filled in box indicates α globin exon sequence, open box indicates 5' or 3' non-coding or intron sequence, hatched box indicates H4 histone coding sequence and stippled box indicates SV40 origin sequences.

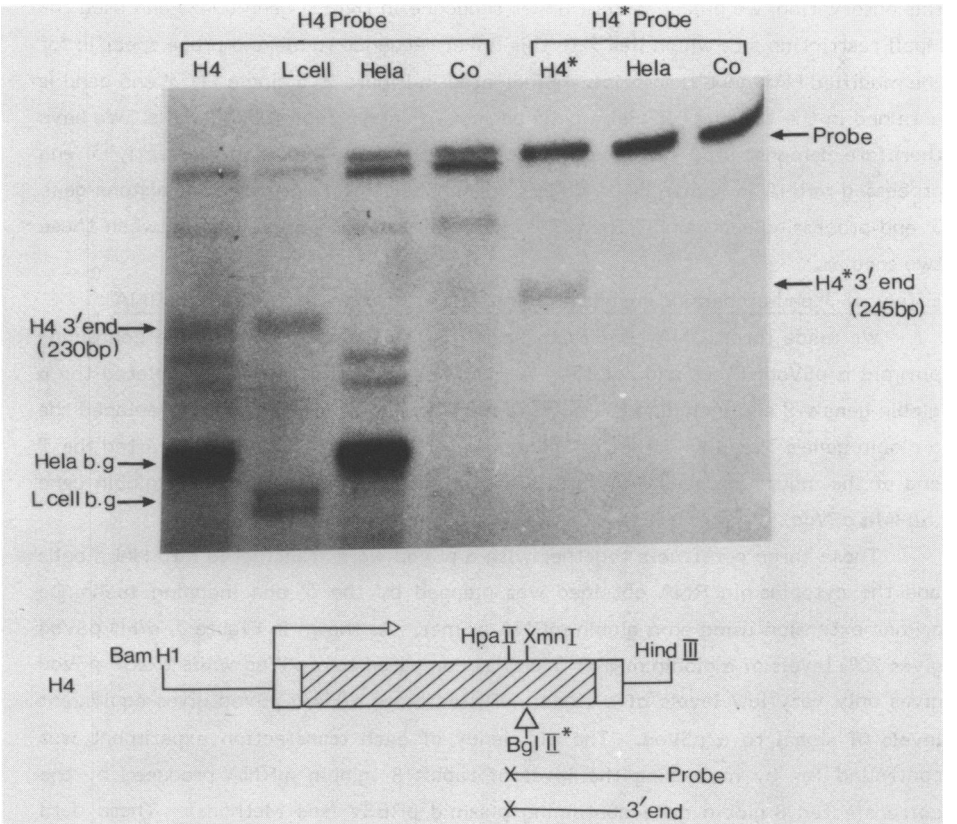

Figure 2. 3' end S1 nuclease analysis. The probe was an end-labelled double-stranded HpaII to HindIII fragment (290 bp for H4 probe and 300 bp for H4* probe). This probe was hybridised to cytoplasmic RNA from HeLa cells transfected with H4 pSVed (H4), from mouse Ltk- cells (L cell) and from HeLa cells which were not transfected (HeLa). The probe hybridized to no RNA and S1 digested, is also shown (Co).

expression of some sea urchin histone genes in heterologous cells results in incorrectly processed mRNA (11). The mouse H4 gene was cloned into the transient expression vector pSVed (8) which contains the SV40 origin of replication and transcriptional enhancer (Fig.1). Following transfection of H4 pSVed into HeLa cells, the RNA obtained was mapped with S1 nuclease using an H4 gene 3' end probe (Fig.2). Both mouse and human HeLa cell RNAs were mapped with the same H4 probe. As indicated the H4 pSVed transfection RNA gives an authentic mouse H4 3' end signal above the background of bands which are due to partial homology of the mouse H4 probe with endogenous human histone mRNAs in the HeLa cells. To further confirm

this observation, we placed a BglII linker sequence in the H4 gene (H4*) and used the HpaII restriction site which lies 5' to this linker sequence to make a probe specific for the modified H4* gene transcript. As indicated in Figure 2, a single H4 3' end band is obtained in the transfected HeLa cells but not in untransfected HeLa cells. We have therefore demonstrated that the mouse histone H4 gene can form correctly 3' end processed mRNA in human HeLa cells. Presumably factors required for histone gene 3' end processing such as U7 SnRNP (12) are sufficiently homologous between these two species.

Effect of 3' end processing signals on the levels of cytoplasmic α-globin mRNA

We made three DNA constructs based on the human α globin gene containing plasmid α pSVod (Fig.1 and ref.13). As shown in Figure 1, we either deleted the α globin gene's 3' end including the poly(A) addition signals (αΔBP pSVod), replaced the α globin gene's 3' end with that of the mouse H4 gene (α/H4 pSVod), or inserted the 3' end of the mouse H4 gene into the third exon of the otherwise intact α globin gene (α/H4/α pSVod).

These three constructs together with α pSVod were transfected into HeLa cells and the cytoplasmic RNA obtained was mapped by the 5' end mapping technique primer extension using an α globin mRNA primer. As shown in Figure 3, α/H4 pSVod gives 20% levels of α globin mRNA 5' end as compared to α pSVod while αΔBP pSVod gives only very low levels of α signal. Furthermore α/H4/α pSVod gives equivalent levels of signal to α pSVod. The efficiency of each transfection experiment was controlled for by measuring the level of rabbit β globin mRNA produced by the cotransfected β globin gene containing plasmid pRBSV (see Methods). These data clearly illlustrate the requirement of 3' end processing signals for the formation of a stable mRNA. The plasmid αΔBP pSVod gives very little stable mRNA while either the presence of the homologous poly(A) site (α pSVod) or histone 3' processing signals (α/H4 pSVod) stabilise the α globin mRNA. However H4 3' processing signals give 5 times less mRNA than with the intact α globin or when the α globin poly(A) site is positioned 3' to the H4 sequence (α/H4/α pSVod).

3' end analysis on α/H4 and α/H4/α pSVod RNA

Having demonstrated that α/H4 pSVod forms stable mRNA based on 5' end analysis, we wished to investigate the 3' ends of this hybrid α/H4 mRNA species. Figure 4A shows S1 nuclease RNA mapping data on α/H4 RNA using a 3' end labelled single strand DNA probe derived from α/H4 pSVod. RNA from cells transfected with α/H4 pSVod gives a band corresponding to authentic H4 mRNA 3' ends together with a full length band due to transcripts reading past the H4 3' end into the pSVod vector sequences. The positions of the H4 mRNA 3' end and HeLa cell histone mRNA background signals are indicated by the L cell and HeLa cell control lanes.

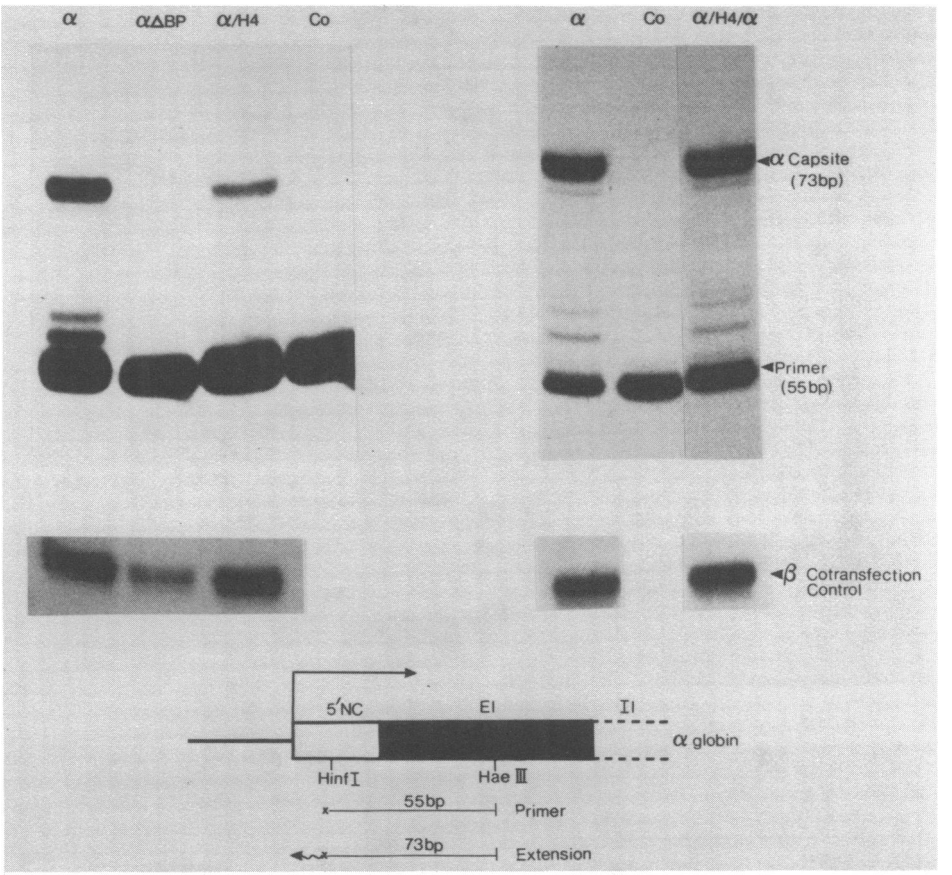

Figure 3. Primer extension 5' end analysis. The probe was a 55 bp end-labelled single-stranded HinfI-HaeIII DNA fragment from α pSVod. The probe was used to analyse cytoplasmic RNA from HeLa cells transfected with α pSVod (α), αΔBP pSVod (αΔBP), α/H4 pSVod (α/H4), α H4/α pSVod (α/H4/α). The primer hybridised to no RNA is also shown (Co). The level of β transcripts from the cotransfected plasmid pRBSV is shown after S1 analysis using the plasmid linearized at the EcoRI site and end-labelled (see Fig.5B).

Approximately 50% of the α/H4 mRNA utilises the H4 3' processing signals while 50% reads through the H4 sequence and possibly utilizes cryptic poly(A) sites in the pSVod vector sequence. Figure 4B shows an equivalent 3' end analysis on α/H4/α RNA versus α/H4 RNA using a probe derived from α/H4/α pSVod. A predominant band is obtained at 440 bp corresponding to the α globin gene poly(A) site. No band was obtained at the authentic H4 3' end position although an artefact band was observed just below this position. To prove that this anomalous band is due to the S1 cutting at

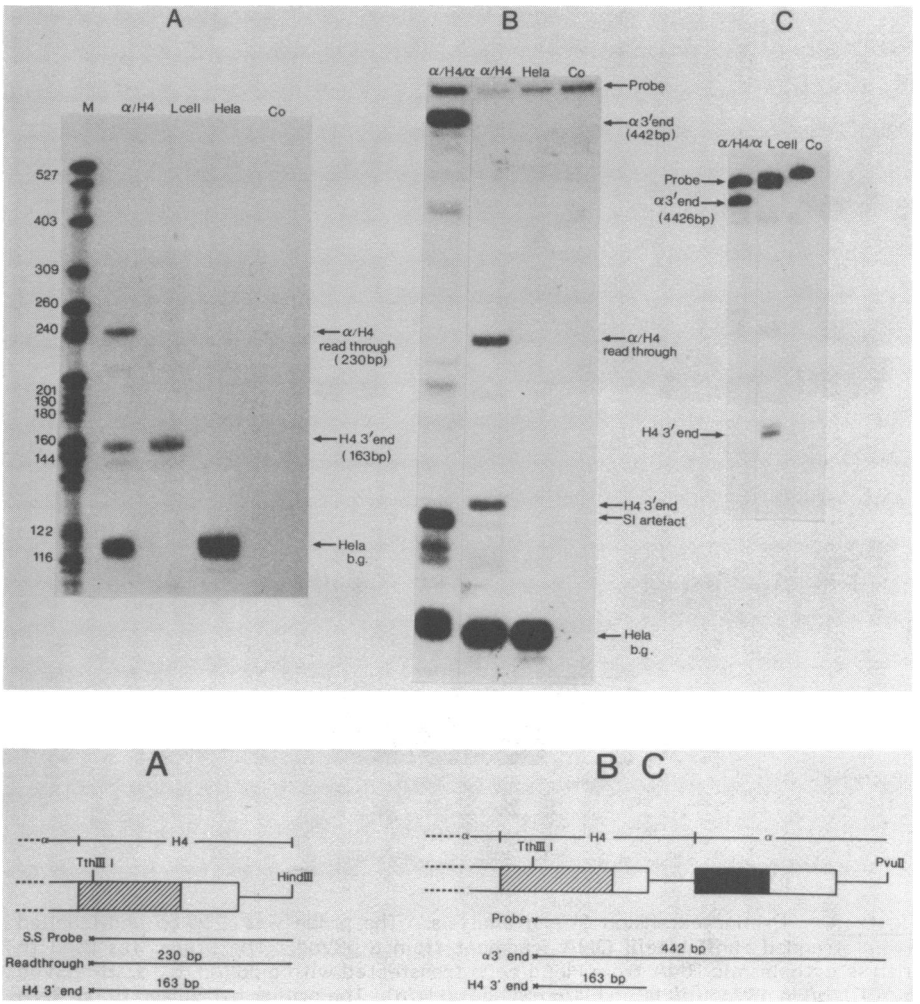

Figure 4.

A. 3' end S1 nuclease analysis. The probe was a single-stranded end-labelled TthIII1 to HindIII fragment from α/H4 pSVod. The probe was annealed to cytoplasmic RNA from HeLa cells transfected with α/H4 pSVod (α/H4), from mouse Ltk⁻ cells (L cell) and from untransfected HeLa cells (HeLa). The probe annealed to no RNA and S1 digested is also shown (Co). HeLa b.g. denotes HeLa background.

B. 3' end S1 nuclease analysis. The probe was a double-stranded end labelled TthIII1 to PvuII fragment from α/H4/α pSVod. The probe was annealed to cytoplasmic RNA fom HeLa cells transfected with α/H4/α pSVod (α/H4/α), α/H4 pSVod (α/H4), nothing (HeLa). The probe annealed to no RNA and S1 digested is also shown (Co).

C. 3' end exonuclease VII analysis. The probe was the same as that used in B, a double-stranded end-labelled TthIII1 to PvuII fragment from α/H4/α pSVod. The probe was annealed to cytoplasmic RNA from HeLa cells transfected with α/H4α pSVod (α/H4/α), cytoplasmic RNA from mouse Ltk⁻ cells (L cell). The probe annealed to no RNA and digested with exonuclease VII is also shown (Co).

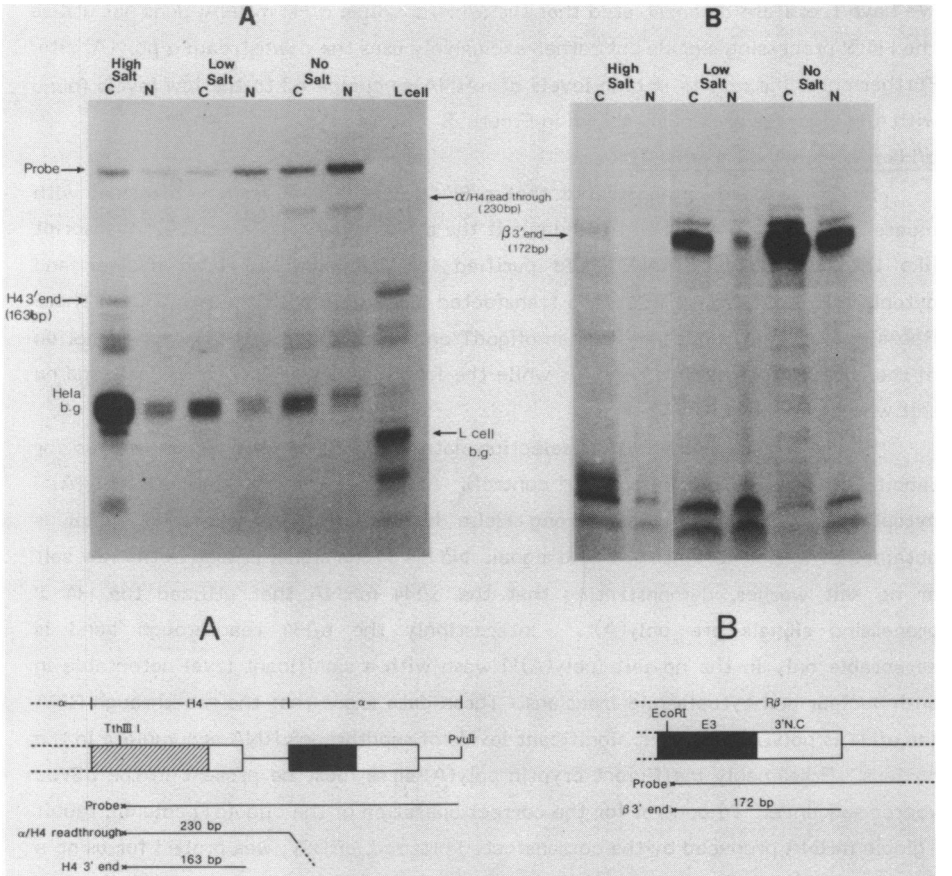

Figure 5. S1 nuclease analysis.

A. The probe was a double-stranded end-labelled TthIII1 to HindIII fragment from α/H4/α pSVod as used in Figure 4B and 4C. The probe was used to analyse cytoplasmic (C) and nuclear (N) RNA from HeLa cells transfected with α/H4 pSVod. The RNA was poly(A) selected using an oligo(dT) column. The poly(A)$^-$ RNA is unbound under high salt (High salt). The poly(A)$^+$ RNA comes off with low salt (Low Salt) and no salt (No Salt) washes. The probe was also hybridized to the cytoplasmic RNA from mouse Ltk$^-$ cells (L cell).

B. The probe was made by linearizing pRBSV with EcoR1 and end-labelling. This probe was annealed to the same RNA samples as described in A.

an internal position in the RNA : DNA duplex (presumably due to AT rich sequence), we carried out exonuclease VII digestion on α/H4/α RNA annealed to the same DNA probe to determine which products of the S1 reaction are due to authentic 3' termini.

As shown in Figure 4C, only the α poly(A) site band in α/H4/α is detectable, proving that the anomalous band at 160 bp does not derive from a true RNA 3' end.

We have therefore demonstrated that the α/H4/α, unlike α/H4 mRNA, does not utilise the H4 3' processing signals but rather exclusively uses the downstream α poly(A) site. Furthermore this results in high levels of mRNA as compared to the low levels found with α/H4, as was previously shown in Figure 3.

α/H4 pSVod mRNA is poly(A)$^-$

Since we have demonstrated that α/H4 mRNA has 3' termini identical with mouse H4 mRNA, we wished to find out if the α/H4 mRNA was a poly(A)$^-$ transcript like the normal H4 mRNA. We purified large amounts of both nuclear and cytoplasmic RNA from HeLa cells transfected with α/H4 pSVod plus pRBSV. These RNAs were poly(A) selected using an oligodT cellulose column. The unbound fraction of the column (high salt) is poly(A)$^-$ while the fractions released with low salt and no salt washes are poly(A)$^+$.

Figure 5 shows the poly(A) selection data for α/H4 mRNA (A) as well as for rabbit β globin mRNA (B) (a poly(A)$^+$ control). As indicated in the high salt (poly(A)$^-$) cytoplasmic fraction, a very strong HeLa histone mRNA background signal is obtained as well as a faint H4 3' end signal. No H4 3' end signal is seen in the low salt or no salt washes, demonstrating that the α/H4 mRNA that utilized the H4 3' processing signals are poly(A)$^-$. Interestingly the α/H4 readthrough band is detectable only in the no-salt (poly(A)$^+$) wash with a significant level detectable in both nuclear and cytoplasmic fractions. These data argue that the readthrough RNA for α/H4 is poly(A)$^+$ but that significant levels of readthrough RNA accumulate in the nucleus. Presumably inefficient cryptic poly(A) sites must be present in the pSVod vector sequence. To control for the correct operation of the oligo(dT) column, rabbit β globin mRNA produced by the cotransfected plasmid, pRBSV, was probed for using a 3' end probe (Fig. 5B). As indicated, β globin mRNA 3' ends were detected in the low salt and no-salt fractions with the most predominant signal in the no-salt cytoplasmic fraction. These data confirm the successful poly(A)$^+$ selection obtained using this oligo(dT) column and therefore demonstrate that the α/H4 mRNA is indeed poly(A)$^-$.

DISCUSSION

The results presented in this study demonstrate that polII transcription initiating off an α globin gene promoter can utilize a histone gene 3' end processing signal and thereby form a stable cytoplasmic poly(A)$^-$ mRNA. This implies that the processing and stabilization of histone genes is independent of the histone promoter and independent of sequences within the 5' coding sequence of the histone gene. It is interesting that experiments in which the 3' processing signals from the snRNA (U1 and U2) genes have been placed at the 3' end of a normally polyadenylated polII gene (β globin promoter) have shown that 3' processing in this case is promoter-dependent

(14). In other words, U1 and U2 3' end processing signals, which like the histone genes do not involve polyadenylation, do not work if they are placed after a β globin gene promoter and will only work if placed 3' to a U1 and U2 promoter. The fact that only 20% of the normal levels of α globin mRNA are obtained when the H4 processing signals are operating suggests that either these transcripts are less stable than normal poly(A)$^+$ transcripts or the histone 3' processing signal is less efficient than the α globin 3' processing signal. When both the H4 3' processing signals and the α poly(A) site are placed in tandem, only the α poly(A) site is utilized, implying that when the two are placed in direct competition with each other, the α globin 3' processing signals are more efficient than the histone 3' processing signal. It is important to note that when two identical and therefore equally efficient poly(A) signals are placed in tandem at the end of a globin gene, the most 5' signal is utilized (15).

It is interesting to consider how much mouse H4 histone sequence is required to confer 3' end processing. We have shown that 250 bp containing the hairpin and 60 bp of 3' flanking sequence are sufficient. This agrees well with the sequences shown to be necessary for 3' end processing of the sea-urchin H2A histone gene (16). It was shown recently that 463 bp containing the 3' half of the mouse H4 histone gene confers 3' end processing if placed in front of an SV40 promoter (17). We have extended these observations by showing that 250 bp are in fact sufficient. In the Luscher et al studies (17), an SV40 poly(A) site was present downstream of the H4 3' end processing signals and yet significant H4 3' end processing still occurred. Because a larger region of H4 3' flanking sequence was present in the SV40/H4 gene construct than in the α/H4/α pSVod construct, it seems possible that transcriptional termination signals are present in the former construct which would prevent the preferential utilization of the downstream SV40 poly(A) site.

In conclusion, these studies clearly demonstrate that histone 3' end processing signals generate stable cytoplasmic poly(A)$^-$ mRNA species from different types of RNA polymerase II genes and that these signals are less efficient than the processing signals present at a normal poly(A) site. Whether this difference in the efficiencies of the two signals has any relevance in vivo is unclear. It will be interesting to define more precisely which sequences confer stability on poly(A)$^-$ mRNA and to understand the molecular basis of this stabilisation process.

ACKNOWLEDGEMENTS

We are grateful to Professor M.L. Birnstiel for the mouse histone H4 gene. This work was supported by a Medical Research Council project grant to NJP, no. G8412674CB.

REFERENCES

1. Lewin B., Genes II, Wiley & Son, 1985
2. Darnell J.E., Nature 297, 365-371 (1982)
3. Nevins J.R., Ann.Rev.Biochem. 52, 441-467 (1983)
4. Mattaj I.W. & De Robertis E.M., Cell 40, 111-118 (1985)
5. Hentschel C.C. & Birnstiel M.L., Cell 25, 301-313 (1981)
6. Whitelaw E. & Proudfoot N.J., Nucl.Acids Res. 211, 7717-7733 (1983)
7. Seiler-Tuyns A. & Birnstiel M.L., J.Mol.Biol. 151, 607-626 (1981)
8. Proudfoot N.J., Rutherford T.R. & Partington G.A., EMBO J. 3, 1533-1540 (1984)
9. Grosveld G.C., de Boer E., Shewmaker C.K. and Flavell R.A., Nature 295, 120-126 (1982)
10. Maniatis T., Fritsch E. and Sambrook J., Molecular Cloning: a Laboratory Manual (1982)
11. Bendig M.M. & Hentschel C.C., Nucl.Acids Res. 11, 2337-2346 (1983)
12. Birnstiel M.L., Busslinger M. & Strub K., Cell 41, 349-359 (1985)
13. Mellon P., Parker V., Gluzman Y. & Maniatis T., Cell 27, 279-288 (1981)
14. Hernandez N. and Weiner A.M., Abstracts from RNA Processing Meeting, Cold Spring Harbor Laboratory (1986)
15. Gil A. & Proudfoot N.J., Nature 312, 473-474 (1984)
16. Birchmeier C., Folk W. and Birnstiel M.L., Cell 35, 433-440 (1983)
17. Luscher B., Stauber C., Schindler R. and Schumperli D., Proc.Natl.Acad.Sci. USA 82, 4389 (1985)