Volume 17 Number 4 1989 Nucleic Acids Research

Transcription from the intron-containing chicken histone $H2A._F$ gene is not S-phase regulated

S.Dalton*, A.J.Robins, R.P.Harvey[1] and J.R.E.Wells

Department of Biochemistry, University of Adelaide, Adelaide, SA 5001 and [1]Walter and Eliza Hall Institute, Royal Melbourne Hospital, Parkville, Victoria 3050, Australia

Received August 23, 1988; Revised and Accepted January 23, 1989 EMBL accession no. X13894

ABSTRACT

The nucleotide sequence of an 8.2 kb *Bam*HI fragment containing the entire chicken histone $H2A_F$ gene has been determined. Unlike the majority of histone genes, the coding region is interrupted by four intervening sequences. While sequencing the 8.2 kb *Bam*HI fragment it was found that the promoter and first exon of an unidentified non-histone gene lies immediately downstream of the $H2A_F$ gene. Studies of $H2A_F$ gene transcription show that, unlike the major core and H1 histone genes, it is not coupled to DNA synthesis.

INTRODUCTION

We have previously identified a cDNA clone (1) from a chick embryonic cDNA library that codes for an extremely variant H2A protein ($H2A_F$). $H2A_F$ is 40% divergent from the most abundant H2A protein in chicken erythroid chromatin and from the predicted amino acid sequence, it seems likely that the $H2A_F$ gene codes for the chicken analogue of the H2AZ histone protein previously characterised in mammalian cells (2,3,4). The H2AZ protein is also related to a histone variant (hv1) characterised in the protozoan, *Tetrahymena*, which is found exclusively in the transcriptionally-active macronucleus (5,6). The ubiquitous nature of histones and the general conservation of their primary structure across wide species barriers suggests that they play a fundamental role in determining chromatin structure. Even so, the existence of histone variants such as $H2A_F$, the enrichment of sub-types in actively-transcribed chromatin, and the occurrence of precisely timed stage-specific variants during differentiation and development suggests they play more than a passive role in nucleosome formation.

In general, histone gene expression is tightly coupled to DNA synthesis in the cell cycle. Genes coding for the major core and H1 histone proteins are expressed predominantly during S-phase (replication-dependent) and this is reflected by an increase of 10-50 fold in levels of their steady-state mRNAs at this time (7,8,9). The accumulation of histone transcripts during S-phase is controlled by the rate of gene transcription and differential transcript stability (9,10,11) culminating in elevated histone protein synthesis coincident with DNA replication. However, there exists a sub-class of replication-independent histone proteins, the synthesis of which is not linked to the cell cycle. For example, accumulation of the tissue-specific linker histone H5

continues during erythroid maturation, long after cessation of cell division. We have previously shown that this can be accounted for by the pattern of H5 gene transcription which is independent of the cell cycle (12).

Here we present the entire genomic sequence of the chicken histone $H2A_F$ gene. This gene not only codes for an extreme H2A protein variant, but its organisation and structure is quite different from other clustered core and H1 genes previously reported in the chicken genome (13). By analysis of steady-state mRNA levels and rates of gene transcription in synchronised cell populations, we show that expression of the $H2A_F$ gene is not temporally modulated during the cell cycle.

MATERIALS AND METHODS

Suppliers

Aphidicolin was a generous gift from the Natural Products Division, National Cancer Institute (Bethesda, Maryland). All radioactive precursors were obtained from Bresatec Ltd., except [^{3}H]thymidine which was purchased from New England Nuclear Corp. Sources of all recombinant DNA clones have been described previously (12).

Fragment Subcloning and Sequencing Strategy

The insert of the $H2A_F$ cDNA clone (1) was used to detect and isolate the chromosomal gene from a chicken genomic library. An 8.2 kb *Bam*HI fragment containing the entire $H2A_F$ gene was isolated from the genomic clone (designated CH2A.F-001), and subcloned into pUC8. A library of random fragments of the 8.2 kb insert was generated in M13mp8 and sequenced by the dideoxy chain termination method. Both strands of the cloned DNA were fully sequenced. Briefly, the insert was concatamerised by ligation and sonicated under conditions chosen to give an average fragment length of 700 bp. The sheared DNA was blunt-ended using Klenow fragment of *E. coli* DNA polymerase and size fractionated on low gelling temperature agarose. DNA in the size range 0.5-1.5 kb was isolated and subcloned into *Sma* 1-digested M13mp8. A library of about 300 recombinants was generated in this manner and sequenced. The computer programmes of Staden (14) were used to align overlaps.

Cell Culture and Synchronisation

Cells used in this study were kindly provided by Dr. T. Graf (Institute for Viral Research, Heidelberg, Federal Republic of Germany) and have been described by us previously (12). All chicken cell lines were grown in Dulbecco modified Eagle Medium (GIBCO Laboratories) supplemented with 10% fetal calf serum (Flow Laboratories) and 2% chicken serum (Commonwealth Serum Laboratories, Australia). Synchronisation of AEV *ts*34 cells (approximately 2 x 10^6/ml) at the G1/S boundary was achieved by the addition of aphidicolin (5 μg/ml in dimethyl sulfoxide) for 20h. Cells were released from the cell cycle block by washing three times in fresh media (minus serum) at 37°C and were resuspended (2 x 10^6 per ml) in fresh medium plus serum supplements. Relative rates of DNA synthesis were determined by pulse

labeling cells with [^{3}H]thymidine according to the method described previously (12).

Isolation and Analysis of Transcripts

Cytoplasmic RNA was isolated from AEV *ts*34 cells by the NP40 method (12). Enrichment of poly(A)+ mRNA species was performed by oligo-dT chromatography (15). Primer extension analysis (16) used to determine the 5' terminus of $H2A_F$ mRNA was performed by annealing a complimentary synthetic gene-specific oligonucleotide (5'-CCAGCCTTCCCACCTGCCATGGTGCCGC-3') to RNA and extending the primer to the 5' end of $H2A_F$ mRNA with reverse transcriptase. Steady-state levels of histone mRNAs were analysed by standard Northern blot hybridisation techniques using gene-specific [^{32}P]-labeled probes (12).

Nuclear Run-on Transcription Assays

Nuclei were isolated as described previously (12) except that cells were lysed by homogenisation in a Dounce homogeniser using a tight fitting pestle. Freshly prepared nuclei (2.0×10^8/ml) were pulse labeled with [^{32}P]UTP (100 μCi) for 15 min at 25 °C in a 150 μl mix consisting of 50 mM Tris hydrochloride (pH 7.9), 5 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM dithiothreitol, 1 mM ATP, 0.5m GTP, 0.5 mM CTP and 0.15 M KCl. The reaction was terminated by the addition of sodium dodecyl sulfate to 0.5% ($^w/_v$) and adjusted to 100 mM NaOAc (pH5.2), 1 mM EDTA. The mixture was phenol-chloroform extracted, brought to 0.2 M NaOAc (pH 5.2), and the RNA was ethanol precipitated. All DNA filter hybridisations involving labeled transcripts were performed as described previously (12).

RESULTS

The H2AF gene contains intervening sequences

Unlike most other histone genes, the $H2A_F$ gene contains intervening sequences (IVS, see Fig. 1). The first IVS is 1536 bp long and occurs directly after the initiation codon; the second is 246 bp long and occurs between codons 26 and 27; the third is 1370 bp long and occurs between codons 64 and 65; and the fourth is 2681 bp long and occurs within codon 108. The third and fourth IVS are each, to a large extent, made up of simple repeating sequences. Repeats from one IVS are not related to those in another, but each class of repeat is represented many times in the chicken genome as determined by Southern blot experiments (data not shown).

Mapping 5' terminus of H2AF mRNA

To determine the 5' terminus of chicken $H2A_F$ mRNA, primer extension analysis was performed on poly(A)+ cytoplasmic RNA isolated from AEV *ts*34 cells with an $H2A_F$ gene specific oligonucleotide (28-mer). This generated an extension product of 144 bases corresponding to the designated 'cap' site located at an A residue, 124 bp upstream from the 'ATG' codon (Fig. 1B and 2).

Promoter sequences in the $H2A_F$ gene

There are several discernible sequences in the $H2A_F$ promoter region which are likely to

B

GATCCGCAGGCCACGCCTCCGCCCCCCCACCCCGCCTTCAGCCTCTGCGTGCCGTGTGCAAAGGAAAAACAGCCGTTAAACGGGCCGCGGGGCGGGGCCAGCGCGGCGGCCAACAAGCG 121

GCGCGGCGATTGGCGGCAGCGGGGCGTTCAAATCAGTACGGGGCGCTCCGATTGGCTCGGGGTCGCCGGGGCAGGCCGGGCTCCCTCCTCGCGCTCGGGCGGGTATTGTCTGGCGGCTCCT 242

GAGGCGGTTCGAGCGGCGGATTCGGGGACCGGGGCTCGGGCGGCGGCGCACC ATG GTGAGGAGCGGGGCCGGGGGGTGAGGAGACGGTCTGAGGGGAGCGAGGACCGCGCTGGGGGC 361

met

TGAAGGGAGCGGGGCCCAAACCGTGGGAGCGTTGCGCTCTCTGTGGGCGGGAAGGCCTGAGGGAGAGAGGAGAAACGTCACTGCTGGGGAACCGGCGGGCAGAGCCGCCACAGAACGGCG 482

GGGTTGGAAGGGGCCTTTAGGAGCGTCTCATCCCAACCCTCTGCTATAGGCGGGGATACCGCCCGGCAGAGCAGCGCTCGGCGCCCCACCCGGCCTGAGCGCTTCCAGGGAGGGGGAACGG 603

CTGCAGCCCTCAGCGGGGAATGGCAGAGCGGCCGAAAGGCCCCGACCCGCTGTTTGGTGCTGCCGGGGGGGCTGCAGGTCCTGCCCTGTGCTCCTTCCCGTGCCCCTGTCCTGCATTCCCT 724

GTGCCCTTTCCCTGTATTCGTGCCGTGTATCCCTGCCCCGTGCCCCTGTCGTGTGCTCCTGCCCTGTGCCAGGAGGGCAGTCCCGAATATGGCACTTCCCAATGCTTTCCACCGCTCTCCC 845

ACTTCCACATCCCACACTGCCGTCCCTATGGAGTCTCTAAAGAAGCCCCATTTCAACACCAGCCCCCAAAGCGCCGCTCCTGTATGGGCTGTCAGCTCCACGCTGTGCTGCGCCGCTCCCG 966

AGCTGAATGCTTTCCCATTGCAGCTTTGCTGGTGGGGGTTTTATTCGGCGTTTGGGCTGTGCTGTTCCTAAAGGGTGCTGTGCAGGGATGTGGGTTTTGTGCAGCGATGCGTTCCCCGTGG 1087

GGAGGCAGCGCTGTATTTTAGGGGTTAATGTGCTCAGTTAGGGTTTTGTATGGACCCCTTTCATGACAGCACGTGTTGGCTGTGCTCATAGCATCATAGAATGGCCTGGGTTGGAAAGGAC 1208

CACGATGCTCATCCAGTTCCAACCCCCTGCTATGTGCAGGGTCGCCAACCAGCAGCCCAGGCTGCCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTCCAGGGATGGGGCATCCACAGCC 1329

TCCTTGGGCTGCCTGTTCAGTGCGTCACCACCCTCTGGTTGAAAAACTTCCTCCTCATATCCAACCTAAACCTCCCCTGTCTCATTTTAAAACCATTCCCCCTCATCCTATCACCGTCCAC 1450

CCTCGTAAACAGCCATTGCTCCTCCTGTTTATACGCTCCCTTCAAGTACTGGAAGGCCGCAACGAGGTCXCCCCGCAGCCTTCTCTTCTCCAAGCTAAAGAAGCCCGGTTCCCTCCACCTT 1571

TCCTCACAGGAGAGGTGCTGCAGCCCTCTGACCATCTTAGTGGCCCTCCTCTGGACCCACTCCAAGAGCTCCACGTCCTTCACCTCGCAGGGCTGGTTGTGCTCAGCCCCCTGTGCTGGCT 1692

GCGCTCGGCTCCCAGCTGCGTCCCACCTACATTTGGAACCAGGAGGAGAACAGAGGGAGGGCCCGGCGCGGTGTCTGAGCCTTCTGCGGCCAAAGCACGGCACGGAGCTGCGTGCTGCGCTC 1813

ACGGCACTCTCCCTCCCGCAG GCA GGT GGG AAG GCT GGG AAG GAC AGC GGG AAG GCC AAG GCG AAG GCG GTG TCT CGC TCG CAG AGA GCC GGA TTG 1909

ala gly gly lys ala gly lys asp ser gly lys ala lys ala lys ala val ser arg ser gln arg ala gly leu

CAG GTAACGAGCGGCGGCCGCAGTGCTCCCCATCTGTACCACGCTAATTGCTGTAAGCACCGCTCCCGCTGCTGACACAGTGCCCCGCGACGCCTTTTTCCCCCATTCCAAGGTGAAGTT 2029

gln

GCGCTGCTGTGCGTTGTCCTTCGGTCTCGCTCACCCTCTGGCTGTGCCTGCAGGGCAGGGCAGGGCCGGATGCTCCCGACGTGAGGCCGTGAGCCGCTGCCCGCCCGCTCATCTGCCTCTC 2150

CGTTGCAG TTC CCC GTG GGC CGC ATC CAT CGG CAC CTG AAG ACG CGC ACC ACG AGC CAT GGG CGG GTC GGG GCC ACC GCC GCC GTG TAC AGC 2242

phe pro val gly arg ile his arg his leu lys thr arg thr thr ser his gly arg val gly ala thr ala ala val thr ser

GCT GCC ATC CTC GAG TAT CTC ACT GCT GAG GTGGGXGCCGGGCCTCGGATCTGGGTGACAATGGGGTGGGGGAAATGCTGCTTTGCAAGGGGAGGGGAGGGCTGTCAGATA 2353

ala ala ile leu glu tyr leu thr ala glu

CATCCCAGCCGTGGGCTGTGTGGGACGGCTGTGCTGTGCTGCTGCCCGCTCACCTCGGCAGCACAGCTGAGGGTCTCTGAAGTGTTGCTGGAATGCTGTGAGCCTGGATGCTATGGGGCAG 2474

CACTATGGGATGGTGCTATGGGGCAGTGCTTTGGGGTTCTGCTGTGGGGTGATGCTGTGGGGCAGAGCTATGAGGCAGTGCTTTGGGGTTCCACTATGGGACGGTGCTATGGGGTGATGCT 2595

GTGGGGCAGAGCTATGAGGCAGTGCTTTGGGGTTCCACTATGGGATGACGCTATGGGGTAGTGCTTTGGGGTTCTGCTGTGGGGCAGAGCTATGAGGCAGTGCTTTGGGGTTCTGCTATGG 2716

GGTGATGCTGTGGGGCAGAGCTATGAGGCAGTGCTTTGGGGTTCCACTGGGATGATGCTATGCAGCAGTGCTTTGGGGTTCTGCTATAGGGTGATGCTATGGGGTAGTGCTTTGGGGTTCC 2837

ACTATGAGATGGTGCTATGGAGCAGTGCTTTGGGGTTCTGCTATGGGGTGATGCTGTGGGGCAGAGCTATGAGGCAATGCTTTGGGGTTCTGCTATGGGGTGACACTTTGGGGCAGTGTTT 2958

TGGTCTGGCACTATGGGGAGGCAGTATGGGGTGACACCGTAAGATGACGCTGTGGGGCACAGTTCTGGGGTGACGCTATGGGGCACAGTTTTGCAGTCACACTACAGGGCAGCGCTATGGG 3079

GTGATGCTATGGGGCGACAGTGTGGTGTGATTCTATGGGGCTGTGATACAGGGCAGTGCTGTGGGGTGACACTATGGGTAATGGCACAGGGCAGTGCTATGGGGTAGCACTAAATGGTGG 3200

CTATGGGGCAATGCTTTTGGGCAGCACTGTGGGGCAGTGCTATGGGGTGACAGTATGGGGAACATGGGATGATGCTATGGAGTAACATTATGGGACAATGCTCTGAGGTGACGCTATGGGG 3321

CAGCACTATGGGGTTACTGGGATGGTGCTATGGGGCAGTGCTTTGGGGTTCTGCTATGGGGCAGTGCTGTGGGATGGTGCTATGGGGCCGTGCTTTGGCGTGACGCTGCAGGGTGATGCTG 3442

TGACACAGCGCTATGGGGCAGTGCTTTGGGGTTCCACTATGGGGCAGCACTATGGGATGGTGCTATGGGGCATTGCTTTGGGATGACATTAGGGGGTAATGCTGTGACGCAGCACTTTGGG 3563

GCTCTGCTATGGGGCAGCAGCTCAGCTGAGCGCAGCTTTACCAACGTGAGGCCCCACGCTGTGCCCCCATCCCCCCGCAG GTC CTG GAG TTG GCA GGC AAC GCC TCC AAG 3672

val leu glu leu ala glu asn ala ser lys

GAC CTG AAG GTG AAG CGC ATC ACT CCC CGC CAT TTG CAG CTG GCG ATC CGC GGC GAC GAA GAG TTG GAT TCC CTC ATC AAA GCC ACC ATA 3762

asp leu lys val lys arg ile thr pro arg his leu gln leu ala ile arg gly asp glu glu leu asp ser leu ile lys ala thr ile

GCG GGG GGA G GTAGGTGGCACCCATAGGGGCACACAGTGGCACCCATAGGAGCACCCAGCATTACCTAATGGCACCCACTGGCATCCATAGGGGCACACAGCATTACCTAATGGCAC 3879

ala gly gly g

CACAGGGGCACCCAGCAGTACTTTATGGCACCACAGGGGCACCCAGCAGTACCTAATGGCACCCATAGAGGCACACAGTGGCACCCATAGGAGCACCCAGCGGTACCTTATGGCACCCACT 4000

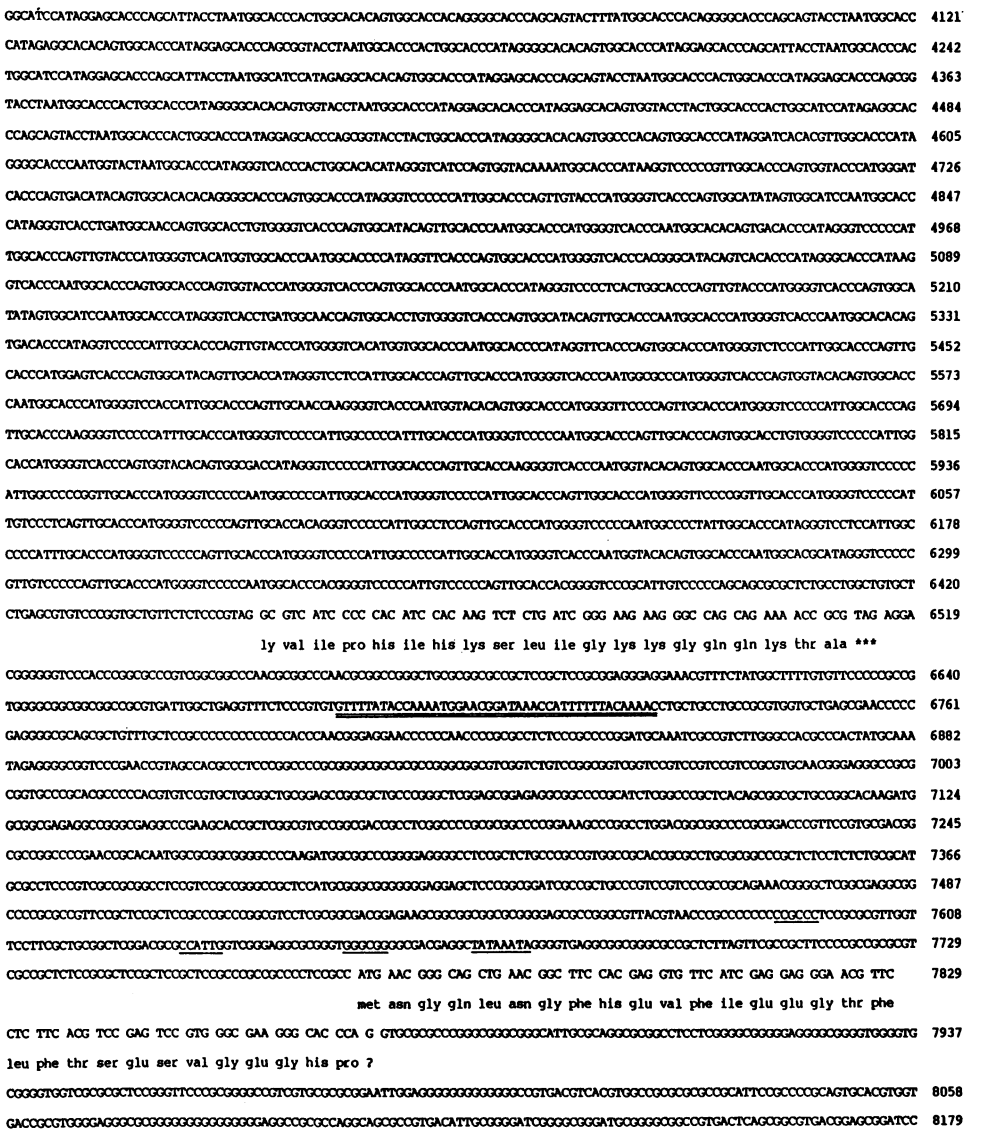

Figure 1:
Organisation and complete genomic sequence of the chicken $H2A_F$ gene. (A) Relative positions of introns (solid line) and exons for $H2A_F$ (filled boxes) and the downstream gene (open box) are shown along the 8.2 kb *Bam*HI fragment. (B) The genomic sequence of an 8.2 kb *Bam*HI fragment containing the entire $H2A_F$ and partial sequence of the downstream gene are shown. Positions of putative control elements are indicated (underlined) in addition to a dyad symmetry element 3' to the $H2A_F$ gene (double underline).

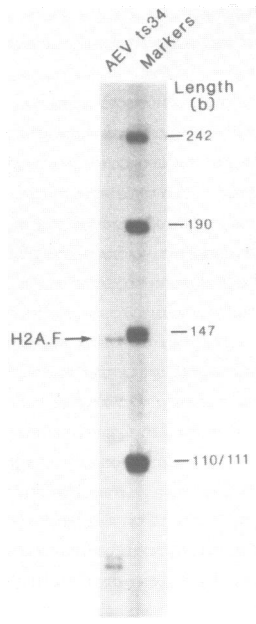

Figure 2:
5' End Mapping of the $H2A_F$ Transcript. Primer extension analysis of AEV *ts*34 poly(A)+ RNA (300 ng) was performed using a synthetic oligonucleotide complimentary to $H2A_F$ RNA which was extended using reverse transcriptase. A specific extension product of 144 bases is shown and from this the predicted cap site is indicated (arrow) in Fig. 1B.

play a role in the expression of the gene (see Fig. 1B). Firstly, 20 bp upstream from the cap site is an unusual TATA-box sequence (5'-TTCAAA-3'), which does not contain the highly conserved tri-nucleotide 'ATA'. While this is somewhat unusual, there are other examples of RNA polymerase II-transcribed genes with atypical TATA-boxes which function adequately *in vivo* (17). Secondly, the $H2A_F$ 5' region contains two motifs initially described by McKnight (18), as the proximal and distal G/C elements in the Herpes Simplex Virus tk (HSV tk) gene promoter which bind the transcription factor Sp1 and are required for the optimal transcription of this gene. These two motifs (5'-GGGCGG-3' and 5'-CCGCCC-3') are positioned 50 and 121 bp upstream from the 'TATA-box' respectively in the $H2A_F$ promoter, similar to their relative position in the HSV tk gene. It has not yet been determined which of these elements are functional in promoter activity.

3' $H2A_F$ gene sequences

Unlike most histone mRNAs, the $H2A_F$ transcript is polyadenylated. Previously (1), we were unable to determine the exact location of the 3' end of the mRNA as the cDNA characterised did not extend back to the poly (A) tail. Another cDNA clone which contained the 3' end of the gene was isolated from a chicken liver cDNA library and sequenced (A. Robins,

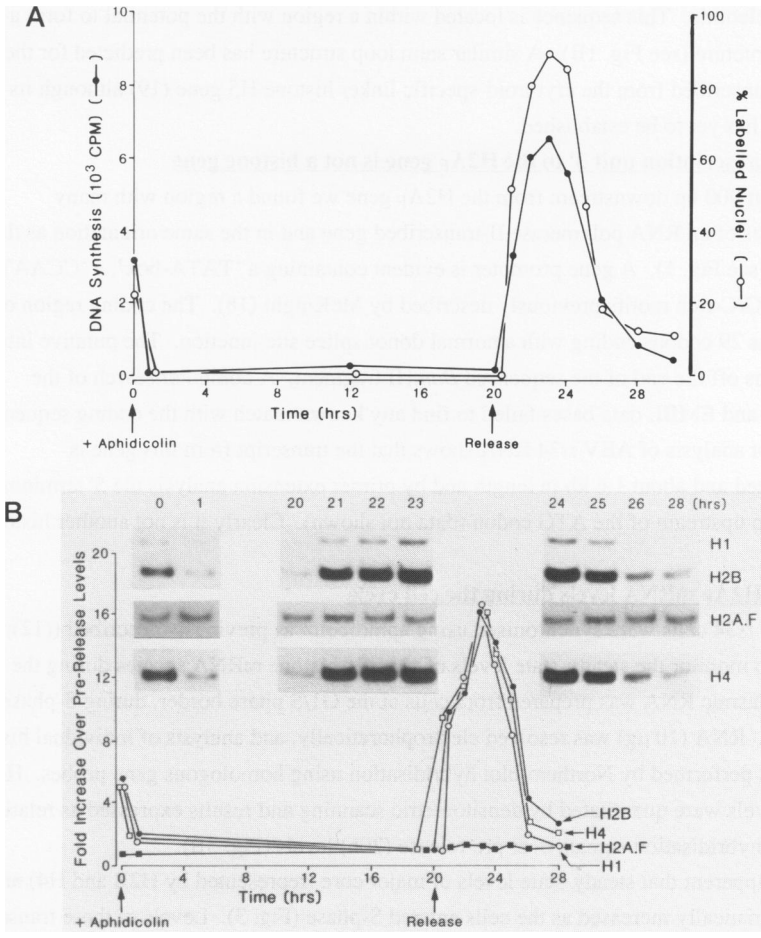

Figure 3:
(A) Kinetics of DNA Synthesis in AEV *ts*34 cells. Where indicated, aphidicolin (5 μg/μl) was added to randomly growing cultures of cells and was removed by washing in fresh media after 20 h. Cells were pulse labelled with [^{3}H]thymidine for 15 min. The rate of radioactivity incorporation into trichloroacetic acid-insoluble material, indicating the rate of DNA synthesis, and the percentage of labelled nuclei, indicating the percentage of cells in S-phase, are shown. (B) Steady-state Levels of Individual Histone mRNAs During the Cell Cycle. Cytoplasmic RNA samples (10 μg) prepared at the times indicated were resolved on 1.5% agarose gels after denaturation and levels of histone mRNAs were detected by Northern blot hybridisation with specific gene probes. Densitometric quantitation of the autoradiograms shown in each panel is also represented. Symbols: ○, H1 mRNA; ●, H2B mRNA; ■, $H2A_F$ mRNA; and □, H4 mRNA. Data is representative of the mean values from two independent experiments.

unpublished data). This data showed that the $H2A_F$ mRNA does not possess an 'AATAAA' sequence found near the 3' end of most genes which produce polyadenylated transcripts. Instead, the sequence 'GATAAA' was identified which differs from the canonical consensus sequence by

only one nucleotide. This sequence is located within a region with the potential to form a stable stem loop structure (see Fig. 1B). A similar stem loop structure has been predicted for the 3' end of mRNA transcribed from the erythroid-specific linker histone H5 gene (19) although its significance has yet to be established.

A nearby transcription unit 3' to the $H2A_F$ gene is not a histone gene

About 900 bp downstream from the $H2A_F$ gene we found a region with many characteristics of an RNA polymerase II-transcribed gene and in the same orientation as the $H2A_F$ gene (see Fig. 1). A gene promoter is evident containing a 'TATA-box', a 'CCAAT-box' and the two G/C-rich motifs previously described by McKnight (18). The coding region of the gene contains 29 codons ending with a normal donor splice site junction. The putative intron sequence runs off the end of the sequenced *Bam*HI fragment. A computer search of the GENBANK and EMBL data bases failed to find any known match with the coding sequence. Northern blot analysis of AEV *ts*34 RNA shows that the transcript from this gene is polyadenylated and about 1.6 kb in length and by primer extension analysis the 5' terminus is located 72 bp upstream of the ATG codon (data not shown). Clearly it is not another histone gene.

Analysis of $H2A_F$ mRNA levels during the cell cycle

AEV *ts*34 cells were synchronised using aphidicolin as previously described ((12); see Fig. 3A). To monitor the steady-state levels of specific histone mRNA species during the cell cycle, cytoplasmic RNA was prepared from cells at the G1/S phase border, during S-phase, and in G2-phase. RNA (10 μg) was resolved electrophoretically, and analysis of individual histone mRNAs was performed by Northern blot hybridisation using homologous gene probes. Histone transcript levels were quantitated by densitometric scanning and results expressed as relative increases in hybridisation signals over pre-release (20h) levels (Fig. 3B).

It is apparent that steady-state levels of major core (represented by H2B and H4) and H1 mRNAs dramatically increased as the cells entered S-phase (Fig. 3). Levels of these transcripts are maximal approximately 3 to 4 hours into S-phase. Accumulation of H2B mRNA during this period (Fig. 3B) is typical of the pattern of histone transcripts generated from other replication-dependent genes. During mid-S-phase, the concentration of H2B mRNA was elevated 17-fold compared with levels at the G1/S phase border or after completion of S-phase (Fig. 3B). Following the peak in steady-state transcript accumulation, levels declined sharply in parallel with the rate of DNA replication, as monitored by the kinetics of [^{3}H]thymidine incorporation (Fig. 3A).

Identical RNA samples were probed to detect $H2A_F$ transcripts (Fig. 3B). In contrast to the major core and H1 mRNAs, levels of $H2A_F$ transcripts were essentially invariant at all cell cycle time points tested and certainly displayed no coupling with DNA replication. This pattern of $H2A_F$ gene expression is not due to low level expression of a sub-population of non-dividing, spontaneously-differentiating cells, since *in situ* hybridisation experiments have shown that the

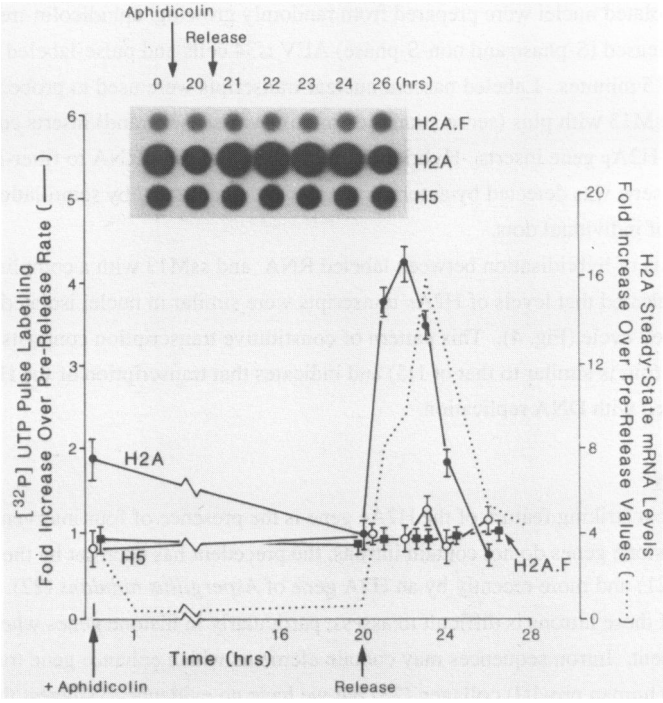

Figure 4:
Transcription of the $H2A_F$ Gene Throughout the Cell Cycle. Nuclei were isolated from aphidicolin-blocked and aphidicolin-released AEV *ts*34 cells and pulse-labeled as described (Ref. 12). Labeled RNA obtained from nuclei isolated at various stages of the cell cycle was used to probe DNA dots (5 μg). Hybridisation of RNA to DNA dots was determined by autoradiography and quantitated by scintillation spectroscopy. Each bar represents +/- the standard error of the mean from three separate experiments. Each hybridisation was performed in duplicate. The dotted line represents the profile of steady-state core histone mRNA levels (represented by H2A; see Ref. 12) superimposed directly on the figure without adjustment to the y-axis.

$H2A_F$ gene is expressed in >95% of AEV *ts*34 cells (20). Therefore, the apparent constitutive pattern of $H2A_F$ gene expression was contributed by dividing cells in which the majority of core and H1 genes are expressed periodically.

Transcription of the $H2A_F$ gene

It is known that enhanced rates of transcription during DNA replication make a major contribution to the accumulation of histone mRNA from S-phase regulated genes (9,10,11). As $H2A_F$ mRNA is polyadenylated it may be much more stable than its normal cell cycle regulated counterparts and therefore might be transcribed only during S-phase, but remain as a stable entity throughout the cell cycle. To test this possibility, run-on transcription experiments were

performed. Isolated nuclei were prepared from randomly growing, aphidicolin-treated, and aphidicolin-released (S-phase and non-S-phase) AEV *ts*34 cells and pulse-labeled with [^{32}P]UTP for 15 minutes. Labeled nascent nuclear transcripts were used to probe filter-immobilised ssM13 with plus (sense strand) or minus (non-sense strand) inserts corresponding to H2A, H5, and $H2A_F$ gene inserts. Hybridisation of labeled nuclear RNA to filter-immobilised ssM13 with inserts was detected by autoradiography and quantitated by scintillation spectrometry of individual dots.

Analysis of hybridisation between labeled RNA and ssM13 with a complimentary $H2A_F$ gene insert indicated that levels of $H2A_F$ transcripts were similar in nuclei isolated at different stages of the cell cycle (Fig. 4). This pattern of constitutive transcription contrasts with that of the H2A gene (but is similar to that of H5) and indicates that transcription of the $H2A_F$ gene is not co-ordinated with DNA replication.

DISCUSSION

The most striking feature of the $H2A_F$ gene is the presence of four intervening sequences. While most histone genes do not contain introns, the precedent has been set by the chicken H3.3 variant gene (21) and more recently by an H2A gene of *Aspergillus nidulans* (22). The significance of these introns is difficult to assess, particularly in histone genes where clasically introns are absent. Intron sequences may contain elements which enhance gene transcription as, for example in human pro-1(I) collagen (23) but we have no evidence to suggest that this occurs with $H2A_F$. We do know that the promoter and leader sequences are sufficient to promote transcription of a heterologous gene in the *Xenopus* oocyte system (data not shown).

The three chicken replication-independent histone gene types characterised so far (H5, H3.3, and $H2A_F$) share a number of other properties. Amongst these three however, the H5 gene is unique as it contains no introns (19). All three have polyadenylated mRNAs, which is unusual for vertebrate histone genes. The H3.3 genes appear to utilise an 'AUUAAA' signal (24) rather than 'AAUAAA' for processing and polyadenylation while the 'GAUAAA' sequence probably serves a similar role in $H2A_F$ mRNA 3' end formation. For the H5 gene, neither of the above hexamer sequences is present in the 3' end of the transcript but we have previously predicted a possible secondary loop structure near the 3' polyadenylation site which may contribute to 3' processing in this case (19).

cDNA clones encoding variant H2A proteins have recently been isolated from human, cow and rat (25), *Drosophila* (26), sea urchin (27) and *Tetrahymena* (28). These H2A variants are translated from polyadenylated transcripts and, with the exception of sea urchin, are encoded by single copy genes.

In common with the H5 gene and the H3.3 genes in chicken, $H2A_F$ is not closely linked to other core and H1 genes in the chicken genome. Neither is the expression of these variant genes S-phase regulated. The biological role of replication-independent histones is not clear,

although the ability of H5 to displace H1 histones and to condense chromatin may be linked to repression of the chicken genome in erythroid cells. If $H2A_F$ is functionally similar to hv1 of *Tetrahymena* (6), then it is possible that it will be associated with transcriptionally-active chromatin. Further studies are in progress to investigate the distribution and biological role of the $H2A_F$ protein.

ACKNOWLEDGEMENTS

We would like to thank Paul Krieg for critical reading of this manuscript. This work was supported by a Commonwealth Special Grant to the Department of Biochemistry, University of Adelaide. Stephen Dalton was the recipient of a CPGR Award.

*To whom correspondence should be addressed

REFERENCES

1. Harvey, R.P., Whiting, J.A., Coles, L.S., Krieg, P.A. and Wells, J.R.E. (1983). Proc. Natl. Acad. Sci. USA *80*, 2819-2823.
2. Ball, D.J., Slaughter, C.A., Hensley, P. and Garrard, W.T. (1983). FEBS Letts. *154*, 166-170.
3. Hatch, C.L., Bonner, W.M. and Moudriankakis, E.N. (1983). Science *221*, 468-470.
4. West, M.H.P. and Bonner, W.M. (1983). Comp. Biochem. Physiol. *76*, 455-464.
5. Allis, C.D., Glover, C.V.C., Bowen, J.K. and Gorovsky, M.A. (1980). Cell *20*, 609-617.
6. Allis, C.D., Richman, R., Gorovsky, M.A. Ziegler, Y.S., Touchstone, B., Bradley, W.A. and Cook, R.G. (1986). J. Biol. Chem. *261*, 1941-1948.
7. Alterman, R.-B., Ganguly, S., Schulze, D.H., Marzluff, W.F., Schildkraut, C.L. and Skoultchi, A.I. (1984). Mol. Cell. Biol. *4*, 123-132.
8. DeLisle, A.J., Graves, R.A., Marzluff, W.F. and Johnson, L.F. (1983). Mol. Cell Biol. *3*, 1920-1929.
9. Heintz, N., Sive, H.L. and Roeder, R.G. (1983). Mol. Cell. Biol. *3*, 539-550.
10. Graves, R.A. and Marzluff, W.F. (1984). Mol. Cell. Biol. *4*, 351-357.
11. Sittman, D.B., Graves, R.A. and Marzluff, W.F. (1983). Proc. Natl. Acad. Sci. USA *80*, 1849-1853.
12. Dalton, S., Coleman, J.R. and Wells, J.R.E. (1986). Mol. Cell. Biol. *6*, 601-606.
13. D'Andrea, R.J., Coles, L.S., Lesnikowski, L., Tabe, L. and Wells, J.R.E. (1985). Mol. Cell. Biol. *5*, 3108-3115.
14. Staden, R. (1980). Nucl. Acids Res. *8*, 3673-3694.
15. Aviv, H. and Leder, P. (1972). Proc. Natl. Acad. Sci. USA *69*, 1408-1412.
16. McKnight, S.L. (1982). Cell *31*, 355-365.
17. Darnell, J.E. (1982). Science *297*, 365-371.
18. McKnight, S.L., Kingsbury, R., Spence, A. and Smith, M. (1984). Cell *37*, 253-262.
19. Krieg, P.A., Robins, A.J., Colman, A. and Wells, J.R.E. (1982). Nucl. Acids. Res. *10*, 6777-6785.
20. Dalton, S. (1987). PhD Thesis, University of Adelaide, Adelaide, South Australia.
21. Engel, J.D., Sugarman, B.J. and Dodgson, J.B. (1982). Nature *297*, 434-436.
22. May, G.S. and Morris, N.R. (1987). Gene *58*, 59-56.
23 Roussow, C.M.S., Vergeer, W.P., du Plooy, S.J., Bernard, M.P., Ramirez, F. and de Wet, W.J. (1987). J. Biol. Chem. *262*, 15151-15157.
24. Dodgson, J.B., Yamamoto, M. and Engel, J.D. (1987). Nucl. Acids Res. *15*, 6294.
25. Hatch, C.L. and Bonner, W.M. (1988). Nucl. Acids Res. *16*, 1113-1124.

26. van Daal, A., White, E.M., Gorovsky, M.A. and Elgin, S.C.R. (1988). Nucl. Acids Res. *16*, 7487-7497.
27. Ernst, S.G., Miller, H., Brenner, C.A., Nocente-McGrath, C., Francis, S. and McIsaac, R. (1987). Nucl. Acids Res. *15*, 4629-4644.
28. White, E.M., Shapiro, D.L., Allis, C.D. and Gorovsky, M.A. (1988). Nucl. Acids Res. *16*, 179-198