Characterization of a cDNA clone coding for a sea urchin histone H2A variant related to the H2A.F/Z histone protein in vertebrates

Susan G.Ernst*, Heidi Miller, Carol A.Brenner¹, Catherine Nocente-McGrath, Susan Francis and Robert McIsaac

Department of Biology, Tufts University, Medford, MA 02155, USA

Received February 17, 1987; Revised and Accepted May 15, 1987

ABSTRACT

A cDNA clone coding for a sea urchin histone H2A variant has been isolated. The coding region of the clone has been sequenced and the sequence found to be closely related to the H2A.F sequence in chickens. The nucleotide sequence of the sea urchin H2A.F/Z is 74% conserved when compared to chicken H2A.F and 51% conserved compared to sea urchin H2A early and 60% compared to sea urchin H2A late. The nucleotide-derived amino acid comparisons show that H2A.F/Z is 97% homologous with H2A.F in chickens and 57% and 56% homologous when compared to sea urchin H2A early and late respectively. There are between 3-6 copies of the H2A.F/Z sequence in the *S. purpuratus* genome. The H2A.F/Z gene sequence codes for the previously identified H2A.Z protein. All embryonic stages and adult tissues tested contain mRNA for H2A.F/Z. The mRNA appears in the poly A⁺ RNA fraction after chromatography over oligo dT cellulose.

INTRODUCTION

Most eukaryotic organisms utilize multiple non-allelic histone variants (reviewed 1-4). Investigation of several recently discovered variants has revealed that the genes for some of these minor histone variants exhibit novel features for histone genes. Unusual features of these variant histone genes include: introns, polyadenylation of the 3' end of the mRNAs, and replication - independent expression of these genes (reviewed 3 and 5,6). Although unusual for histone genes, each of these features is a common characteristic for non-histone genes. In addition to non-allelic histone variants similar to those observed in other systems, sea urchins have three developmental classes of histone genes that are expressed at sequential and overlapping times during embryogenesis and in the adult (7). As yet, developmental class switching of histone genes has not been observed in other organisms.

Histones play a fundamental role in organizing DNA into nucleosomes and, together with other nuclear proteins, into higher orders of chromatin

structure (reviewed 8,9). In view of the basic structural role played by histones, it is not at all clear why organisms utilize multiple histone variants. Two obvious exceptions are in sperm and in avian erythrocytes, where variant histones appear to be characteristic of altered chromatin structure. Indeed, in both of these instances, at least one of the 5 histone proteins is replaced by a variant isoform (3).

Another histone variant that may characterize structurally altered chromatin is H2A.F/Z. Several lines of evidence, from a number of different systems, suggest that the histone variant H2A.F/Z may be associated with genes active in transcription (10-13). We report here isolation of a sea urchin cDNA clone coding for H2A.F/Z. In this study we have sequenced the clone, looked at its representation within the genome and investigated the developmental expression of the H2A.F/Z gene.

MATERIALS AND METHODS

Animals and embryos.

Procedures for maintaining Strongylocentrotus purpuratus and spawning, fertilization and growth of embryos were as previously described (14).

Isolation of λ 108.

A cDNA clone, λ 108, was isolated in a differential screen of a gastrula λ gt10 cDNA library (gift of Drs. T. Thomas and E. Davidson), using 32 P-labeled cDNA from RNAs of primary mesenchyme and presumptive endoderm/ectoderm cells as hybridization probes. The λ 108 sequence was originally identified for further study because it appeared slightly enriched in presumptive endoderm/ectoderm cells, compared to primary mesenchyme. More importantly for this study, λ 108 represents a sequence complementary to mRNA that peaks in abundance at the mesenchyme blastula stage. Phage DNA was isolated (15) and the sea urchin sequence, a 1.55 kb fragment, was liberated from the vector by digestion with EcoRI.

Subcloning and restriction mapping.

The sea urchin λ 108 sequence was subcloned into pUC8 (16). Restriction mapping was carried out using the isolated 1.55 kb sea urchin fragment.

Sequencing.

pSp108 was sequenced using both chemical sequencing (17) and the dideoxy chain termination technique (18). The Protein Identification Resource, National Biomedical Research Foundation at Georgetown Medical Center was used to search for a computer match of the sequence.

Genomic Southernns and copy number.

Aliquots of *S. purpuratus* sperm DNA from a single animal were digested with three restriction enzymes and the DNA fragments separated by electrophoresis on a 0.8% agarose gel. DNA was transferred to nitrocellulose for hybridization with ³²P-labeled pSp108. Hybridization was carried out overnight at 65°C in 3 X SET, 10 X Denhardt's solution and 0.1% SDS. To establish the number of copies of the H2A.F/Z gene in the genome, total genomic sea urchin DNA from a single animal and the 1.55 kb sea urchin fragment of pSp108 equivalent to 0.2, 1, 2, and 5 copies were spotted onto nitrocellulose paper. The filter was hybridized to nick-translated, ³²P-labeled pSp108 as above. After exposure of the autoradiograph, hybridization dots were cut from the filter and quantitated by liquid scintillation counting.

RNA and poly A⁺ RNA isolations.

Total RNAs were isolated from eggs, embryos and adult structures by modification of a procedure of Chirgwin *et al.* (19) employing guanidine hydrochloride in the lysis buffer. The RNA was pelleted through a cushion of 5.7M CsCl. Total RNA from mesenchyme blastula stage embryos was fractionated into poly A⁺ and poly A⁻ fractions by passage over oligo dT cellulose (20).

RNA blot analysis.

RNA samples were denatured, separated by electrophoresis on 1% agarose formaldehyde gels and transferred to nitrocellulose paper as described (21). Hybridization to nick-translated, ³²P-labeled pSp108 was overnight at 65°C in 0.32M phosphate buffer pH 6.8, 0.8% Denhardt's solution, 0.08% SDS and 10% dextran sulfate. In order to quantitate the relative abundance of RNA, hybridization bands were cut from the nitrocellulose filter by alignment with the autoradiograph. The bands were then counted by liquid scintillation counting.

RESULTS

Isolation of λ clone 108.

The cDNA clone, λ 108 was screened from a gastrula λ gt10 library, using cDNA probes from blastula stage embryos. Due to the high relative abundance of mRNA complementary to λ 108 at the mesenchyme blastula stage, compared to a decrease in message abundance at gastrula, we reasoned that λ 108 may represent a gene sequence important in early development and undertook a detailed study of this clone.

Figure 1: Restriction map of pSp108.

The λ 108 clone codes for a variant H2A histone gene, H2A.F.

The sea urchin cDNA fragment was cut out of λ gt10 using EcoRI and the fragment subcloned into the plasmid pUC8. A restriction map of this subclone, called pSp108, is shown in Figure 1. The entire coding sequence and a portion of the 3' untranslated region of pSp108 have been sequenced. The longest sequenced open reading frame is 379 nucleotides in the orientation shown in Figure 1. Since the clone is approximately 1.55 kb in length, the 3' untranslated region of the mRNA is at least 1.1 kb long. However, it is likely that the 3' untranslated end is longer than 1.1 kb, as the mRNA is approximately 1.7 kb and no poly A tail was found at the 3' end of the clone. Sequence data were used to search the nucleotide data bank of the Protein Identification Resource, National Biomedical Research Foundation. A match was found with a chicken histone variant, H2A.F (22).

Interestingly, the first search, carried out when only 50 nucleotides had been sequenced identified the chicken H2A.F variant, but not the sea urchin early H2A histone gene sequence which was also in the data bank (23). Complete sequence data and subsequent analysis revealed the reason for the initial identification of the chicken variant and not the sea urchin sequence. Figure 2 compares the nucleotide sequences of the histone H2A.F coding regions from chicken and pSp108. Sequence data are available for the entire coding region of the chicken sequence (22). Assuming that the sea urchin protein is the same size as the chicken H2A.F, we have all but 8 coding nucleotides at the 5' end of the sea urchin cDNA clone. Out of 379 coding nucleotides 99 are mismatches, resulting in 74% homology between the genes of these evolutionarily distant organisms.

When the predicted amino acid sequences of these genes are compared, the evolutionary conservation is even more striking. Of 125 amino acids, there are only four amino acid changes or a 3% difference between the

Figure 2: Comparison of sea urchin H2A.F sequence to chicken H2A.F. The nucleotide sequence for clone pSp108 is compared to the published nucleotide sequence for chicken H2A.F (22). Mismatches are indicated by a vertical line.

See Urchin H2A.F/Z gly lys ala gly lys asp ser gly lys ala lys ala lys
Chicken H2A.F ala gly gly arg his ala gly lys asp ser gly lys ala lys ala lys
Mammals H2A.Z lys ala gly lys asp ser gly lys ala lys thr lys

ala val ser arg ser ala arg ala gly leu gln phe pro val gly arg ile his arg his
ala val ser arg ser gln arg ala gly leu gln phe pro val gly arg₂ ile his arg his
ala val ser arg ser gln arg ala gly leu gln phe pro val gly₁ arg₂

leu lys thr arg thr thr ser his gly arg val gly ala thr ala ala val thr ser ala
leu lys thr arg thr thr ser his gly arg val gly ala thr ala ala val thr ser ala

ala ile leu glu tyr leu thr ala glu val leu glu leu ala gly asn ala ser lys asp
ala ile leu glu tyr leu thr ala glu val leu glu leu ala gly asn ala ser lys asp

leu lys val lys arg ile thr pro arg his leu gln leu ala ile arg gly asp glu glu
leu lys val lys arg ile thr pro arg his leu gln leu ala ile arg₂ gly asp glu glu
his leu gln leu ala ile arg₂

leu asp ser leu ile lys ala thr ile ala gly gly gly val ile pro his ile his lys
leu asp ser leu ile lys ala thr ile ala gly gly gly val ile pro his ile his lys

ser leu ile gly lys lys gly ser gln lys ala thr term
ser leu ile gly lys lys gly gln gln lys thr ala term

Figure 3: Amino acid sequence of sea urchin H2A.F/Z predicted from DNA compared to the predicted sequence in chicken and protein sequence in mammals. The predicted amino acid sequences for the sea urchin H2A.F/Z and chicken H2A.F are compared. Mismatches are indicated by a line over the amino acid. These sequences are compared to peptide data from calf (24) and mouse (25) with mismatches indicated by a line under the amino acid.

Figure 4: Nucleotide comparison of sea urchin H2A.F/Z with sea urchin H2A.F early. Sequence data for H2A early in sea urchins (23) compared to the H2A.F/Z sequence with mismatches marked by vertical line. The * indicates a missing nucleotide compared to the other sequence.

conclude that pSp108 codes for the sea urchin equivalent of H2A.F.

Amino acid sequence data exists for a fragment of a calf H2A variant called H2A.Z (24) and for 2 fragments of mouse H2A.Z (25). By aligning the mammalian H2A.Z sequences with the chicken and sea urchin H2A.F sequences, the available data suggest that H2A.F is equivalent to the mammalian H2A.Z variant (Figure 3 and 13,24,25). Further, the mammalian H2A.Z protein variant has been shown to comigrate on 2-dimensional gels with sea urchin H2A.Z and the two proteins have seven arginine and eight lysine tryptic peptides in common (26). The Z variant in sea urchins has also been called Y6 and M (7).

Sea Urchin H2A.F/Z is significantly divergent from other sea urchin H2A genes. Comparison of the sea urchin variant with other sea urchin H2A genes shows significant divergence among the members of the sea urchin H2A gene family. In order to compare the sequences, they were first aligned using the amino acids and this positioning was used for the nucleotide comparisons. To match the amino acid sequences between H2A early and H2A.F, 2 deletions and one addition have been made in the H2A early sequence. The assignment of one amino acid deletion near the N-terminal end of the H2A early protein sequence has been made after comparison to the H2A late sequence (27). Also H2A.F has one additional C- terminal amino acid. Figure 4 illustrates that between H2A early (23) and H2A.F there are

```

Sea urchin H2A.F/Z      gly lys ala gly lys asp ser gly lys ala lys ala lys
Sea urchin H2A early    ser gly arg ala lys * ser gly lys ala arg thr lys

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

leu lys thr arg thr thr ser his gly arg val gly ala thr ala ala val thr ser ala
leu arg lys gly asn tyr ala lys * arg val gly gly gly ala pro val thr met ala

ala ile leu glu tyr leu thr ala glu val leu glu leu ala gly asn ala ser lys asp
ala val leu glu tyr leu thr ala glu ile leu glu leu ala gly asn ala ala arg asp

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

leu asn ser leu ile lys ala * thr ile ala gly gly gly val ile pro his ile his
leu asn lys leu leu gly gly thr val ile ala gln gly gly val leu pro asn ile gln

lys ser leu ile gly lys lys gly ser gln lys ala thr tem
ala val leu leu pro lys lys thr ala lys ser ser tem

```

Figure 5: Amino acid comparison between sea urchin H2A.F/Z and H2A early. Amino acid sequences derived from nucleotide data are compared for these two sea urchin H2A genes. The * indicates a missing amino acid in the corresponding sequence. The nine boxed amino acids are the H2A conserved sequence found in all H2A genes examined (28).

184 mismatched nucleotides out of a total of 379, or 49% divergence between the two sequences. A similar comparison between H2A late and H2A.F/Z (27) reveals only a 40% divergence in the nucleotide sequence (data not shown). Although the H2A.F/Z sequence is significantly divergent from either H2A early or late, the conservation of the nine amino acids of the H2A invariant peptide (20) (boxed in figure 5) demonstrates that H2A.F/Z is an authentic H2A sequence. Interestingly, while the degree of conservation on a nucleotide level is higher between H2A late and H2A.F (60%) compared to H2A early and H2A.F (51%), the conservation on an amino acid sequence level is about the same: 56% and 57% respectively.

Representation of H2A.F/Z within the *S. purpuratus* genome.

To investigate the representation of pSp108 within the genome, the fragment was labeled with ^{32}P by nick-translation and hybridized to total genomic DNA. Aliquots of DNA from sperm of an individual animal were digested with three different restriction enzymes. The restriction fragments were separated on an agarose gel, transferred to a nitrocellulose filter, and hybridized to the ^{32}P labeled pSp108. Results of the hybridization show a pattern consistent with a low copy number of this sequence per genome (Figure 6A).

A determination of the number of copies of H2A.F/Z was carried out for comparison with copy numbers for H2A early and H2A late genes. The early histone genes, which are transcribed following fertilization and throughout

Figure 6: Representation of the H2A.F/Z gene within the *S. purpuratus* genome. A. Ten μ g samples of sea urchin total genomic DNA from a single individual were digested with 1) Hind III, 2) Bgl II, 3) EcoRI, 4) EcoRI and resolved on a 1% agarose gel. Hybridization was with H2A.F/Z DNA labeled with 32 P by nick-translation in 3x SET, 10x Denhardt's solution, and 0.1% SDS at 65°C. (Molecular weight markers are given in kb.) B. Reconstruction dot blot analysis of copy number. Sea urchin total genomic DNA from a single animal, position 1 and 2, and pSp108 DNA equivalent to a) 0.2, b) 1, c) 2 and d) 5 copies were dotted onto a nitrocellulose filter. The filter was hybridized as in A. C. Quantitation of copy number. Hybridization dots were cut from the filter in B and quantitated by liquid scintillation counting. Counts were averaged and plotted; ● represents pSp108 standards and ▲ represents genomic DNA.

cleavage, are present approximately 400 times per haploid genome (23). The late genes, those that start transcription towards the end of cleavage, are represented at 5-12 copies per genome (29,30). The number of copies of the

Figure 7: The H2A.F/Z mRNA is restricted to the poly A⁺ RNA fraction. Mesenchyme blastula RNA was separated into poly A⁻ and poly A⁺ fractions by passage over oligo dT cellulose. RNAs were separated on a 1% agarose formaldehyde gel and hybridized to ³²P-labeled pSp108. Five ug of RNA were loaded of non-fractionated RNA lane 1, poly A⁻ RNA lane 2, and poly A⁺ lane 3.

H2A.F/Z sequence in the genome was determined by reconstruction dot blot analysis. The results, shown in Figure 6B and C, demonstrate that the gene for H2A.F/Z is present approximately 4 times per haploid genome. This experiment has been carried out several times and results range from 3-6 copies. Therefore, consistent with the whole genome Southern results, it appears that the histone H2A.F/Z variant gene is present only a few times per haploid genome.

pSp108 mRNA is found in the poly A⁺ RNA fraction.

Since histone mRNAs are generally non-polyadenylated, it was unexpected to have isolated a histone sequence from a cDNA library. To measure the distribution of H2A.F/Z mRNA between the poly A⁺ RNA and poly A⁻ RNA fractions, total cellular RNA from mesenchyme blastula stage embryos was fractionated on an oligo dT cellulose column. Total RNA, poly A⁺ RNA and poly A⁻ RNA were electrophoresed on a formaldehyde agarose gel and transferred to nitrocellulose. The RNAs were hybridized to ³²P-labeled pSp108 DNA. Results, shown in Figure 7, demonstrate that H2A.F/Z mRNA appears to be exclusively in the poly A⁺ fraction and is apparently absent from the poly A⁻ fraction. Preliminary sequence data from the 3' end of pSp108 have not revealed a poly A tail nor a typical polyadenylation site

Figure 8: Developmental expression of H2A.F/Z mRNA. A. Total RNA was isolated from 1) egg 2) 16-cell stage (5.5 hrs) 3) mesenchyme blastula (23 hrs) 4) gastrula (36 hrs) and 5) pluteus stage embryos (72 hrs). RNA (15 ug lane) was separated on a 1% agarose formaldehyde gel under denaturing conditions, transferred to nitrocellulose and the blot hybridized to 32 P-labeled pSp108 DNA. This is a long exposure of the blot to enable visualization of RNA at early cleavage. B. Several RNA blots were run and the bands cut out and counted. Not all stages were run on all gels, but mesenchyme blastula RNA was always included as a standard. All RNAs run on the same gel were from the same batch of embryos. Results are normalized to % of counts in mesenchyme blastula lane and averaged from several blots. Egg represents 3 points; 16 cell stage 3 points; mesenchyme blastula 5 points; gastrula 4 points; pluteus 5 points. Error bars represent the range of values obtained.

AATAAA (31). However, as noted earlier, it is suspected that the 3' end of the clone may not be full length.

H2A.F/Z mRNA is represented in all developmental stages and adult structures. The relative abundance of the mRNA for H2A.F/Z was determined at selected developmental stages. RNAs were separated under denaturing conditions on an agarose formaldehyde gel and transferred to nitrocellulose. Hybridization to 32 P-labeled pSp108 demonstrated that all developmental stages contained a 1.7 kb RNA complementary to pSp108 and that this transcript is most abundant in mesenchyme blastula stage embryos (Figure 8A). Several RNA blots, using RNAs isolated from separate batches of embryos were analyzed by cutting out the bands and quantitating the relative hybridization at the various embryonic stages. Figure 8B is a

Figure 9: The mRNA for H2A.F/Z is found in adult tissue. Fifteen ug of RNA from a) intestine, b) coelomocytes c) ovary d) testis and e) tube feet were loaded on 1% agarose gels. RNA blots were hybridized to 32 P-labeled pSp108. This figure represents a composite of several gels and therefore the relative intensities of each band do not necessarily reflect the relative abundance of the histone H2A.F/Z mRNA in each adult structure. Intestine and coelomocyte RNAs are on the same gel and ovary and testis RNAs are on the same gel.

composite of these results where, in each blot, counts are normalized to % abundance at mesenchyme blastula which is set at 100%. Histone H2A.F/Z mRNA is stored in the egg as a maternal transcript. The mRNA abundance decreases slightly during early cleavage and then increases dramatically by the end of cleavage. The peak of mRNA abundance declines as the cleavage rate slows during gastrulation and is maintained at this reduced level throughout embryogenesis. Comparison of the relative hybridization of 32 P-labeled poly A⁺ RNA to pSp108 and to the 3' untranslated region of the actin gene CYI, whose transcript number has been previously quantitated (32), showed that there are approximately 50,000 copies of the H2A.Z transcript per mesenchyme blastula embryo (data not shown).

The representation of mRNA for H2A.F/Z in several adult structures was also investigated (Figure 9). pSp108 hybridized to a 1.7 kb RNA in all the adult structures tested including: intestine, tube feet, coelomocytes, ovary, and testis. Since Figure 9 is a composite of separate hybridizations of different RNA blots, the hybridization signals are not an accurate representation of the relative abundance of H2A.F/Z mRNA among these adult structures. However, inclusion of mesenchyme blastula RNA on an adult RNA blot suggests that the abundance of histone H2A.F/Z mRNA in

adult tissues is less than one third that seen at the mesenchyme blastula stage (data not shown).

DISCUSSION

Sea urchins exhibit a complex developmental pattern of expression of members of the histone multigene family. Details of the developmental switches from one class to another have been worked out from data supplied by a number of investigators. The pattern that emerges is that during embryogenesis there are three sets of histone genes: cleavage stage, early and late, which are expressed at sequential and overlapping times during development and in the adult organism. A great deal is known about the early and late variants at both the gene and the protein sequence levels. The cleavage stage histones, which have been identified on protein gels, have so far eluded cloning. In addition to these major histone variants, there are also recently discovered sperm-specific late histone H-1 and H-2b subtypes (33). We have isolated a cDNA clone coding for a histone H2A variant gene that exhibits features uncharacteristic of the major classes of histone genes in sea urchins.

Isolation and characterization of a cDNA clone for the sea urchin histone H2A.F/Z gene has supplied information which furthers our understanding of this unusual H2A histone variant and the gene which codes for it. This is the second report of the isolation of a cDNA sequence coding for the H2A.F variant. The other was a study reporting the cloning of the sequence in the chicken (22), an organism evolutionarily very distant from the sea urchin. The conservation at the amino acid level for these two proteins is a remarkable 97%. Additional information in the literature demonstrates that the protein coded by H2A.F in sea urchins and chickens is nearly identical to the H2A.Z protein in mammals and H2A.Z in sea urchins (figure 3, 13,24-26). In recognition of the correlation between the nucleic acid and protein data, we call the sea urchin sequence H2A.F/Z.

Direct comparison of the sea urchin sequence with the sequence in "lower eukaryotes" is limited by available information. Drosophila has a variant H2A protein, previously called D2 and now named H2A.2 (34) that may be partially equivalent to H2A.F/Z, but sufficient sequence data for critical evaluation is so far lacking. In the protozoan Tetrahymena, the hv-1 variant of the H2A protein shows relatedness to both H2A.F/Z and the mammalian H2A.X histone variants. The carboxyl end of hv-1 appears to

share antigenic determinants with H2A.X, while the amino-terminal end of hv-1 has 70% of 91 amino acids in common with H2A.F/Z and only 56% of the same 91 amino acids in common with the Tetrahymena major H2A histone protein (13). A recent and suprising finding demonstrates that a yeast mitochondrial DNA binding protein, called HX has minor but recognizable homologies compared to the chicken H2A.F sequence (35). Predictably, the same homologies obtain when comparing the sea urchin H2A.F/Z sequence to the nucleotide-derived amino acid sequence for HX. However, the significance of these similarities is, at present, obscure.

When comparing sea urchin H2A.F/Z to the other sequenced sea urchin H2A variants, the conservation between H2A.F/Z in the sea urchin and higher organisms becomes even more noteworthy. At the nucleotide level, the sea urchin H2A.F/Z sequence is conserved 51% compared to sea urchin H2A early, 60% compared to sea urchin H2A late and 74% with chicken H2A.F. Comparison at the level of nucleotide-derived amino acid sequence reveals a 57% conservation between H2A.F/Z and H2A early, a 56% conservation with H2A late and 97% conservation with chicken H2A.F. Other classes of sea urchin H2A genes, H2A early and late, are much more similar to each other than to H2A.F/Z. The strong evolutionary conservation across phyla, compared to the significant divergence of H2A.F/Z relative to other H2A sequences of the same species, argues against H2A.F/Z being the result of evolutionary drift of a multigene family. Instead, it suggests a conserved, unique and essential function for this H2A.F/Z variant.

The H2A.F/Z sequence remains distinct from the major H2A genes in a number of important ways. The gene is represented in the genome at a relatively low copy number, 3-6, compared to approximately 400 copies for early and 5-12 copies for each of the late genes (29,30). In addition the mRNA has an exceptionally long 3' untranslated tail, at least twice as large as the next longest sea urchin histone mRNA. The mRNA fractionates in the poly A⁺ RNA fraction from oligo dT cellulose. In sea urchins, as in most other organisms, the major histone mRNAs are nonpolyadenylated, with the exception of H-1 which appears to be represented in both the poly A⁺ and the poly A⁻ fraction (36,37). However, one of the more unusual features of the H2A.F/Z sequence, and perhaps one of the most important when considering its functions, is the expression of the mRNA and the corresponding representation of the protein.

In sea urchins the levels of histone H2A.F/Z found in the chromatin of embryos from early cleavage to gastrulation generally represents 1-10% of

the total H2A (26); a small, but reproducible portion of the total. Studies presented here demonstrate that H2A.F/Z mRNA is represented in all developmental stages and adult tissues tested. Preliminary data from embryonic cell fractionations (38) and in situ hybridization (39) suggest that distribution of H2A.F/Z mRNA is not lineage specific. From mRNA and protein studies, histone H2A.F/Z is the only known histone variant in the sea urchin that is found in every cell type looked at so far. Furthermore, H2A.F/Z appears to consistently represent a minor portion of the total H2A population, no matter which major H2A variant is present. This suggests that H2A.F/Z is serving a conserved function throughout the life of the organism. Studies from other systems suggest that H2A.F/Z, or a closely related protein variant, may be associated with active genes. In mouse, the H2A variant M1 [equivalent to H2A.F/Z (25,26)] was found associated with chromatin that is enriched for transcriptional activity (11). In addition, the hv-1 variant found in Tetrahymena, which is closely related to H2A.F/Z and has some homology to the mammalian H2A.X protein, is found preferentially associated with active genes (13). Our results are consistent with H2A.F/Z being a H2A histone variant that associates with active genes in sea urchins.

ACKNOWLEDGEMENTS

We wish to thank Dr. Rob Maxson for generously sharing his unpublished sequence data on H2A late. Also we thank Drs. T. Thomas and E. Davidson for the gift of the λ gt10 library and Dr. W.R. Crain for actin DNA. We appreciate the thoughtful comments on the manuscript by Dr. Saul Slapikoff and Judy Boyle. We gratefully acknowledge the help of Dr. George Ellmore and Judy Boyle in the preparation of the figures and Valerie Ricciardone for typing the manuscript. This work was supported by a grant from the National Institutes of Health and by Tufts University Faculty Research Awards.

*To whom correspondence should be addressed

¹Present address: Laboratory of Radiobiology, University of California, San Francisco, CA 94143, USA

REFERENCES

1. Hentschel, C.C. and Birnstiel, M.L. (1981) Cell 25:301-313.
2. Maxson, R., Cohn, R., Kedes, L. and Mohun, T. (1983) Ann. Rev. Genet. 17:239-277.
3. OTds, R.W. and Woodland, H.R. (1984) Cell 38:624-626.

4. Stein, G.S., Stein, J.L. and Marzuff, W.F. eds. (1984) Histone Genes: Structure, Organization and Regulation. John Wiley and Sons, New York.
5. Brush, D., Dodgson, J.B., Choi, O.-R., Stevens, P.W. and Engel, J.D. (1985) Mol. and Cell. Biol. **5**:1307-1317.
6. Wells, D. and Kedes, L. (1985) Proc. Natl. Acad. Sci. USA. **82**:2834-2838.
7. Poccia, D. (1986) Internat. Rev. Cytol. **105**:1-65. Bourne, G.H., Jeon, K.W. and Friedlander, M. eds. Academic Press, Orlando, Florida.
8. McGhee, J.D. and Felsenfeld, G. (1980) Ann. Rev. Biochem. **49**:1115-1156.
9. Ettenberg, J.C., Cartwright, I.L., Thomas, G.H. and Elgin, C.R. (1985) Ann. Rev. Genet. **19**:485-536.
10. Allis, C.D., Glover, C.V.C., Bowen, J.K. and Gorovsky, M.A. (1980) Cell **20**:609-617.
11. Gabrielli, F., Hancock, R. and Faber, A.J. (1981) Eur. J. Biochem. **120**:363-369.
12. Allis, C.D., Ziegler, Y.S., Gorovsky, M.A. and Olmsted, J.B. (1982) Cell **31**:131-136.
13. 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.
14. Pittman, D. and Ernst, S.G. (1984) Dev. Biol. **106**:236-242.
15. Kao, F., Hartz, J., Law, M. and Davidson, J. (1982) Proc. Natl. Acad. Sci. USA. **79**:865-869.
16. Vieira, J. and Messing, J. (1982) Gene **19**:259-268.
17. Maxam, A. and Gilbert, W. (1980) Methods in Enzymol. **65**:499-560.
18. Sanger, R., Nicklen, S. and Coulson, A.R. (1977) PNAS (USA) **74**:5463-5467.
19. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochem. **18**:5294-5299.
20. Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA **69**:1408-1412.
21. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA **77**:5201-5205.
22. Harvey, R.P., Whiting, J.A., Coles, L.S., Kreig, P.A. and Wells, J.R.E. (1983) Proc. Natl. Acad. Sci. USA **80**:2819-2823.
23. Sures, I., Lowry, J., Kedes, L.H. (1978) Cell **15**:1033-1044.
24. Ball, D., Slaughter, C.A., Hensley, P. and Garrard, W.T. (1983) FEBS Lett. **154**:166-170.
25. Wu, R.S., West, M.H.P. and Bonner, W.M. (1984) In Stein, G.S., Stein, J.L. and Marzuff, W.F. (eds), Histone Genes Structure Organization and Regulation. John Wiley and Sons, New York, pp 457-483.
26. Wu, R.S., Nishioka, D. and Bonner, W.M. (1982) J. Cell Biol. **93**:426-431.
27. Maxson, R. Personal communication.
28. von Holt, C., Strickland, W.N., Brandt, W.F. and Strickland, M.S. (1979) FEBS Lett. **100**:201-218.
29. Childs, G., Nocente-McGrath, C., Lieber, T., Holt, C. and Knowles, J.A. (1982) Cell **31**:383-393.
30. Maxson, R., Mohun, T., Gormezano, G., Childs, G. and Kedes, L. (1983) Nature **301**:120-125.
31. Proudfoot, N.J. and Brownlee, G. (1976) Nature (London) **263**:211-216.
32. Floyd, E.E., Gong, Z., Brandhorst, B.P. and Klein, W.H. (1986) Dev. Biol. **113**:501-511.
33. Lieber, T., Weisser, K. and Childs, G. (1986) Mol. and Cell. Biol. **6**:2602-2612.
34. Donahue, P.R., Palmer, D.K., Condie, J.M., Sabatini, L.M. and Blumenfeld (1986) Proc. Natl. Acad. Sci. USA **83**:6878-6882.

35. Delahodde, A., Becam, A.M. and Jacq, C. (1986) Nuc. Acid Res. 14:9213-9214.
36. Ruderman, J.V. and Pardue, M.L. (1977) Dev. Biol. 60:48-68.
37. Brandhorst, B.P. (1980) Dev. Biol. 79:139-148.
38. Miller, H. and Ernst, S.G. Unpublished results.
39. Finkle, J. and Ernst, S.G. Unpublished results.