Sequence of cDNAs for mammalian H2A.Z, an evolutionarily diverged but highly conserved basal histone H2A isoprotein species

Christopher L.Hatch and William M.Bonner*

Laboratory of Molecular Pharmacology, Division of Cancer Treatment, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

Received November, 1987; Revised and Accepted January 6, 1988

ABSTRACT

The nucleotide sequences of cDNAs for the evolutionarily diverged but highly conserved basal H2A isoprotein, H2A.Z, have been determined for the rat, cow, and human. As a basal histone, H2A.Z is synthesized throughout the cell cycle at a constant rate, unlinked to DNA replication, and at a much lower rate in quiescent cells. Each of the cDNA isolates encodes the entire H2A.Z polypeptide. The human isolate is about 1.0 kilobases long. It contains a coding region of 387 nucleotides flanked by 106 nucleotides of 5'UTR and 376 nucleotides of 3'UTR, which contains a polyadenylation signal followed by a poly A tail. The bovine and rat cDNAs have 97 and 94% nucleotide positional identity to the human cDNA in the coding region and 98% in the proximal 376 nucleotides of the 3'UTR which includes the polyadenylation signal. A potential stem-forming sequence imbedded in a direct repeat is found centered at 261 nucleotides into the 3'UTR. Each of the cDNA clones could be transcribed and translated *in vitro* to yield H2A.Z protein. The mammalian H2A.Z cDNA coding sequences are approximately 80% similar to those in chicken and 75% to those in sea urchin.

INTRODUCTION

Mammalian cells contain four histone H2A isoproteins named H2A.1, H2A.2, H2A.X and H2A.Z (1). The synthesis of the H2A.1 and H2A.2 isoproteins occurs primarily in concert with DNA replication during the S phase of the cell cycle. The H2A.X and H2A.Z isoproteins are synthesized at a lower but more constant level throughout the cell cycle (2), and, in addition, are synthesized in quiescent cells at 5-10% of the rate found in cycling cells (3). H2A.Z represents about 5% of the total H2A in vertebrates (1). Although evidence has been presented which suggests that H2A.Z may be enriched in transcriptionally active chromatin (4,5), how its function differs from that of the other histone isoprotein species is not yet known.

In previous studies of the H2A isoproteins, H2A.1 and H2A.2 sequences were found to be very similar (6,7) with only a few amino acid substitutions. H2A.X was found to be slightly less similar with amino acid changes in four of eight lysine peptides and one of eight arginine peptides. H2A.Z, on the other

hand, was found to be very different with amino acid changes in seven of seven lysine peptides and seven of nine arginine peptides. However, each of the four H2A isoproteins appeared to be equally capable of complexing with H2B to form H2A:H2B dimers and to be a component of the core histone octamer (8).

While H2A.Z seems to be the H2A isoprotein that is least similar to its counterparts, its sequence also appears to have been highly conserved during evolution (5,9-12). Harvey et al. (11) obtained from chicken the nucleotide sequence of a cDNA encoding what is now recognized to be an H2A.Z homologous protein, H2A.F. More recently, Ernst et al. (12) have obtained almost all of the cDNA coding sequence for another homologous protein, H2A.F/Z, from sea urchin. The first 30 and 32 N-terminal amino acid residues of H2A.Z protein from cow and trout have been sequenced (13,14) and 88 amino acid residues in the N-terminal region of the H2A.Z related Tetrahymena macronuclear histone isoprotein, hvl, have recently been determined (5). Comparison of these polypeptide sequences has shown that the level of conservation is very high in all regions except for the first 27 N-terminal amino acids of the Tetrahymena protein.

In this study, cDNAs from H2A.Z mRNA of three mammalian species (rat, bovine, and human) have been isolated and sequenced. The amino acid sequences of the encoded proteins were found to be identical in the three species. The resulting conservation of the nucleotide sequence extends through the coding region and through several hundred bases of the 3' untranslated region. Each of the human, bovine, and rat cDNAs have been shown by *in vitro* sequential transcription and translation assays to produce a protein which comigrates with mammalian H2A.Z in a two dimensional gel electrophoretic system optimized for the characterization of histone isoproteins.

MATERIALS AND METHODS

Screening of cDNA libraries

A human cDNA library constructed with the polyA⁺ mRNA fraction from HUT78 cells (15) and the Okayama-Berg expression vector (16,17) was a gift from S. Colombini-Hatch (LTCB, NCI, NIH). Rat brain and bovine brain cDNA libraries constructed in the lambda gt10 vector were received at the Cold Spring Harbor Course "Advanced Molecular Cloning" from J. Brosius (18-20). The bovine library was screened with a chicken H2A.F insert kindly provided by J. Wells (11). The human and rat libraries were then screened with an identified bovine H2A.Z insert. The radioactive probes for screening the libraries were produced with a random primer labeling kit from Pharmacia, (21,22) and proces-

sed through an elutip-d column (Schleicher and Schuell). Filters were prehybridized for several hours at 37°C in 50% formamide, 5x SSC, 50 mM sodium phosphate (pH 6.5), 5x Denhardt's solution, 0.04% SDS, 0.4 mg/ml heat denatured herring sperm DNA (Sigma), and 1.5% dextran sulfate, then hybridized overnight in 50% formamide, 5x SSC, 20 mM sodium phosphate, pH 6.5, 1x Denhardt's solution, 0.04% SDS, 0.4 mg/ml heat denatured herring sperm DNA (Sigma), and 7.5% dextran sulfate containing the radioactive probe at a final concentration of 2×10^6 cpm/ml. After hybridization, the filters were first washed under low stringency conditions (2x SSC, 0.1% SDS, 37°C), sealed between plastic and exposed to film. The washing process was repeated at increasing levels of hybridization stringency up to 0.5 x SSC, 0.5% SDS, 65°C. The screens yielded putative H2A.Z inserts of 0.9 and 1.3 kb for the cow, 1.0 kb for the human, and 1.4 and 1.8 kb for the rat.

Sequencing strategy

Each of the putative H2A.Z cDNA inserts from the human, bovine, and rat cDNA libraries was purified and subcloned into the multicloning site of the Bluescript M13 phagemid vector (Stratagene Cloning Systems, San Diego, CA). Partial sequences were determined using exonuclease deletion (23) or by subcloning restricted fragments of each insert. Sanger dideoxy sequencing (24) was performed with denatured double-stranded plasmid DNA using reverse transcriptase and primers specific either for the T3 promoter or for the T7 promoter. The completed sequence of both strands of each H2A.Z insert was obtained using custom synthesized oligonucleotides complementary to specific sequences within the H2A.Z inserts as primers on denatured double-stranded DNA for the SequenaseTM dideoxy sequencing system (USB Corp.).

In vitro transcription and translation of H2A.Z cDNAs

For in vitro transcription, the Bluescript M13 vectors containing the H2A.Z inserts were restricted at a site downstream from the coding sequence, and the appropriate RNA polymerase for the upstream vector promoter used to synthesize run-off transcripts. The RNA produced was treated with RQ-DNase (Promega Biotech) at 40 U/ml for 15 min at 37°C, extracted three times with phenol-chloroform (50:50), then ethanol precipitated in the presence of 0.3 M sodium acetate, pH 5.0. The RNA pellets were rinsed with ethanol, lyophilized to dryness, and resuspended in sterile deionized H₂O. Aliquots were in vitro translated in rabbit reticulocyte lysates or wheat germ extracts (Bethesda Research Labs) lacking leucine and supplemented with L-[4,5-³H(N)]-leucine, 60 Ci/mmol, (New England Nuclear) to a final concentration of 0.16 mCi/ml. After 1 hr at room temperature, the reaction mixtures were treated with 0.09

mg/ml RNase A (Pharmacia) for 15 min at room temperature, and the histones extracted with an equal volume of 1 M HCl, 10 mM BME for 1 hour on ice. The extract was processed for two-dimensional electrophoresis of histone isoproteins as previously described (25). H2A.Z cDNAs from each of the three species were found to support the *in vitro* transcription of mRNAs which could in turn be *in vitro* translated into polypeptides having an identical electrophoretic mobility with mouse H2A.Z.

H	(-106)	CAGTTTGAATCGCGGTGCGACCGAAGGAGTAGGTGCTGGGATCGTCACCGTGGACCGATTG	(-41)
		MetAlaGlyGlyLysAlaGlyLys	[7]
H		GCCTTTTCTCTGCTTGTCTGCTTGAGCTTCAGCGGAATTCGAAATGGCTGGCGGTAAAGCTGGAAAG	(24)
R		(-15) C AC G C G	
B		(0) G	
H		AlaSerGlyLysAlaLysThrLysAlaValSerArgSerGlyAlaGlyLeuGlnPheProValGly	[30]
H		GACTCCGGAAGGCCAAGACAAAGGCGTTTCCGCTCGCAGAGACCGGCTTGCAGTTCACAGTGGG	(96)
R		C C G T G	
B		T G	
H		ArgIleHisArgHisLeuLysSerArgThrThrSerHisGlyArgValGlyAlaThrAlaAlaValTyr	[53]
H		CGTATTCATCGACACCTAAATCTAGGACGACCACTCATGGACGTGTGGGCGGACTGCCGCTGTGTAC	(162)
R		G A C C C A	
B		T C	
H		SerAlaAlaIleLeuGlyTyrLeuThrAlaGlyValLeuGlyLeuAlaGlyAsnAlaSerLysAspLeu	[76]
H		AGCGCAGCCATCTGGAGTACCTCACCGCAGAGGTACTTGAACCTGGCAGGAAATGCATCAAAAGACTTA	(231)
R		GT	
B		T G G	
H		LysValLysArgIleThrProArgHisLeuGlnLeuAlaIleArgGlyAspGlyGlyLeuAspSerLeu	[99]
H		AAGGTAAAGCGTATTACCCCTCGTCACTTCAACTTGTCTGCGGAGATGAAGAATTGGATTCTCTC	(300)
R		C C G C G	
B		C	
H		IleLysAlaThrIleAlaGlyGlyGlyValIleProHisIleHisLysSerLeuIleGlyLysLysGly	[122]
H		ATCAAGGCTACAATTGCTGGTGGTGGTGCATTCACACATCCACAAATCTCTGATTGGGAAGAAAGGA	(369)
R		G C A	
B		A	
H		GlnGlnLysThrValEnd	[127]
H		CAACAGAAGACTGTCTAAAGGATGCTGGATTCTTGTATCTCAGGACTCTAAATACTCTAACAGCTG	(438)
R		^(del A) ^ (add T)	
B		A	
H		TCCAGTGTGGTGAATCCAGTGGACTGTATCTCTGTGAAAAACACAATTTTGCTTTTGTAAATCTAT	(507)
R			
B			
H		TTGAGCAAGTTGGAAATTTAATAGCTTTCCAACCAACCAATTTCTGCATTCGAGTCTTAACCATATT	(576)
R		A	
B		T	
H		TAAGTGTACTGTGGCTTCAAGAGAGCTATTGATTCTGAAGTAGTGGGTTTGTATTGAGTTGACTGTTT	(645)
R		A G	
B		G	
H		TTAAAAAAGTGTGGATTTAATTGTGATGCAGAAATTATAGTAACAAACATTTGGTTTTGTTCAGAC	(714)
R			
B		A A	
H		CTTATTTCCACTCTGGTGGATAAGTTCAATAAAGGTGCATATCCCAAACTAAAA(poly A tail)	
R		C G AAAA(poly A tail)	
B		C GTTGTATATAAATTTGCTT	(783)
H		GATTATAGTAGGAACAGCTTTGTTGAATAGGTATCTTACCTAGCAATAACTTAAGCACATTTCTTCCT	(852)
B		TTAAATTACTGTTAATCTGTCTGTAGATCAACAAAGTTAAAAGGCCCAAGTG	(905)

Figure 1. The nucleotide and amino acid sequences for human, bovine, and rat H2A.Z.

All three cDNAs encode the same amino acid sequence as shown. The sequences of the rat (R) and bovine (B) cDNAs are shown below the human (H) sequence with only the differences shown. The amino acid sequence is numbered [] from the Ala residue, while the nucleotide sequence is numbered () from the preceding start codon. The polyadenylation signal, AATAAA (positions 742-747), is underlined.

RESULTS

Isolation and sequence analysis of the H2A.Z cDNA clones

The complete nucleotide sequences of the H2A.Z cDNAs from human, cow and rat are presented in Figure 1. The cloned human cDNA contains 106 nucleotides of 5'UTR, 387 nucleotides of coding sequence (including start and stop codons), 376 nucleotides of 3'UTR, and 82 residues in the polyA tail. The three regions are 45, 48 and 64% AT respectively. The rat cDNA has 15 nucleotides of 5'UTR, 6 of which differ from the human sequence; 387 nucleotides of coding sequence, 21 of which differ from the human homologue; and 376 nucleotides of 3'UTR followed by a polyA tail. In the rat 3'UTR compared to the human, there is a deletion of the first nucleotide past the stop codon and a possibly compensating insertion 42 nucleotides downstream. The sequence of the bovine cDNA starts with the first nucleotide of the coding sequence and contains 10 differences from the human homologue in those 387 nucleotides. The human, rat, and bovine cDNAs all have a polyadenylation signal, AATAAA (26), at the identical location in the 3'UTR (position 742 in Figure 1). The human and rat cDNAs do contain polyA tails starting at position 763; however, the bovine cDNA sequence continues past position 763 with over 140 more nucleotides apparently without a polyA tail. One notable feature of the 3'UTR of the cDNAs of all three species is the sequence shown in Figure 2. This sequence of 56 nucleotides contains a 19 nucleotide potential stem forming sequence situated between two 18 nucleotide direct repeats with 78% nucleotide positional identity.

The cDNAs of the three mammalian species encode identical polypeptides as the nucleotide changes found simply change codon usage for the same amino acid sequence. The similarities of the coding and 3' untranslated regions among the three mammalian species is shown in Figure 3. It is notable that the

Figure 2. Optimum base pairing scheme for the putative stem-repeat sequence in H2A.Z cDNA (positions 620 to 675).

The conserved repeats are underlined, and lower case is used to indicate base differences in the repeats.

BOVINE	97.2 (98.1)			
RAT	94.3 (98.1)	94.6 (97.9)		
CHICKEN	78.5 (0)	79.3 (0)	79.1 (0)	
SEA URCHIN	75.7 (N.A.)	77.3 (N.A.)	76.0 (N.A.)	74.7 (N.A.)
	HUMAN	BOVINE	RAT	CHICKEN

Figure 3. Similarities among various H2A.Z cDNAs.

The percent similarities of the coding and 3'UTR sequences (in parentheses) in human, bovine, rat, chicken (11) and sea urchin (12) H2A.Z cDNAs are presented. The base insertion and deletion in the rat 3'UTR are considered as a single difference. The first 8 bases of the sea urchin cDNA coding region have not yet been determined and are not included in the calculations. The 5'UTR sequences of human (106 bases) and chicken (70 bases) H2A.Z cDNA were not found to have any significant similarity. (N.A. = not available).

3'UTR is more conserved among the three mammalian species than is the coding region. Comparison of the human, bovine, and rat cDNAs to the homologous cDNA from chicken (H2A.F) and the more recently isolated homologue from sea urchin (H2A.F/Z) shows each of mammalian sequences to be 75-80% similar to both the chicken and sea urchin coding sequences (Figure 3). The chicken and sea urchin coding sequences also show a 75% similarity when compared to each other. None of the mammalian H2A.Z cDNAs exhibit any similarity in the 3'UTR with the chicken cDNA, and the human and rat cDNAs also have no similarity in the 5'UTR with the chicken DNA. This complete lack of conservation of the 3'UTR sequence between mammalian and avian species is in contrast to the recent observation that the proximal 520 nucleotides of the 3'UTR of the human and chicken H3.3 genes have 85% nucleotide positional identity (27,28).

Comparison of known H2A.Z amino acid sequences

The three mammalian H2A.Z polypeptides presented here have identical sequences, but there are differences between the mammalian and more distant species. Table 1 presents these amino acid substitutions for H2A.Z polypep-

Table 1. Amino acid substitutions in H2A.Z homologues

<u>Position</u>	----- Human	-- <u>Species</u> -- Chicken	----- Urchin
14	thr	ala	ala
21	gln	gln	ala
38	ser	thr	asn
123	gln	gln	ser
126	thr	thr	ala
127	val	ala	thr

tides of other species. The overall amino acid similarity is 97% between chicken (11) and mammal and 95% between sea urchin (12) and mammal. A partial N-terminal peptide sequence of trout H2A.Z is identical to the mammalian sequence in 30 of the first 32 amino acids; positions 14 and 29 are ala and gly respectively (14). The previously determined N-terminal amino acid sequence (30 residues) of bovine H2A.Z (13) is identical to the amino acid sequence predicted by these three mammalian H2A.Z cDNA coding sequences. From these comparisons, it appears that the amino acid sequence of the central core of H2A.Z, residues 39 to 122 is very highly conserved and those few changes that have occurred in the amino and carboxy-terminal regions of the protein are rather conservative. Identical migration of the H2A.Z proteins from the different species in two dimensional gel electrophoretic systems (25) indicates that these amino acid changes are without a significant effect on the overall charge, size, or hydrophobicity (Triton X-100 binding) of these proteins.

Comparison of the H2A.Z and H2A.1 amino acid sequences

While the overall similarity between mammalian and urchin H2A.Z is 95%, it is only 59% between mammalian H2A.Z and H2A.1. Between H2A.Z residues 39 to 122, which are identical in mammal and urchin, there are 32 substitutions between mammalian H2A.Z and H2A.1. Figure 4 presents the comparison of these sequences. However, despite an overall similarity of 59%, there are three regions of much higher similarity as noted in Figure 4. One region, spanning residues 22-31, is the longest stretch of identical sequence and contains within it the conserved arginine peptide originally used to help identify H2A.Z as an H2A isoprotein (1). Two of these regions, spanning residues 45-72 and 83-96, are contained in the central core of the H2A.1 molecule, spanning residues 31-95, which have been shown to be crucial for interaction with the H2B molecule to form the H2A:H2B dimer (29). On the other hand, when the hydropathy (30, using IBI-Pustell Software) of the bovine H2A.Z polypeptide was

H2A.1	---	---	ser	gly	arg	gly	lys	gln	gly	gly	lys	ala	arg	ala	lys	ala
H2A.Z	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	1				5				10						15	
H2A.1	lys	thr	arg	ser	42%	100%	-----	-----	100%	23%						
	x	x		x		arg	ala	gly	leu	gln	phe	pro	val	gly	arg	val his
H2A.Z	val	ser	arg	ser	glu	arg	ala	gly	leu	gln	phe	pro	val	gly	arg	ile his
				20					25					30		
H2A.1	arg	leu	leu	arg	lys	gly	asn	tyr	ala	glu	23%	82%	-----	-----		
	x	x	x	x	x	x	x	x	x	x					x	
H2A.Z	arg	his	leu	lys	ser	arg	thr	thr	ser	his	gly	arg	val	gly	ala	thr ala
		35				40						45				50
H2A.1	pro	val	tyr	leu	ala	ala	val	leu	glu	tyr	leu	thr	ala	glu	ile	leu glu
	x			x			x								x	
H2A.Z	ala	val	tyr	ser	ala	ala	ile	leu	glu	tyr	leu	thr	ala	glu	val	leu glu
					55				60					65		
H2A.1	leu	ala	gly	asn	ala	40%	-----	-----	40%	93%	--					
						ala	arg	asp	asn	lys	lys	thr	arg	ile	ile	pro arg
H2A.Z	leu	ala	gly	asn	ala	ser	lys	asp	leu	lys	val	lys	arg	ile	thr	pro arg
				70				75					80			
H2A.1	his	leu	gln	leu	ala	ile	arg	asn	asp	glu	glu	leu	asn	lys	leu	leu gly
							x						x	x	x	x
H2A.Z	his	leu	gln	leu	ala	ile	arg	gly	asp	glu	glu	leu	asp	ser	leu	ile ---
	85					90				95						100
H2A.1	lys	val	thr	ile	ala	gln	gly	gly	val	leu	pro	asn	ile	gln	ala	val leu
	x					x				x		x	x	x	x	
H2A.Z	lys	ala	thr	ile	ala	gly	gly	gly	val	ile	pro	his	ile	his	lys	ser leu
					105				110						115	
H2A.1	leu	pro	lys	lys	thr	glu	ser	his	his	lys	ala	lys	gly	lys		
	x	x			x	x	x	x	x	x	x	x	x	x		
H2A.Z	ile	gly	lys	lys	gly	gln	gln	lys	thr	val	---	---	---	---		
				120				125								

Figure 4. Sequence comparison of the bovine H2A.1 and H2A.Z isoproteins.

The regions of greatest similarity are indicated with dashed lines. Sequence differences are noted with an x. Residue numbers correspond to the H2A.Z sequence.

compared to that of the bovine H2A.1 polypeptide (Figure 5, left and center) the two molecules were found to be quite similar in this regard throughout their lengths, irrespective of the degree of primary sequence positional identity. Significantly different hydropathy plots were observed for the bovine H2B polypeptide (Figure 5, right) and the bovine H3 and H4 polypeptides (not shown).

Figure 5. Hydropathy plots of the histone polypeptide sequences.

Sequences analyzed are bovine H2A.Z (left), bovine H2A.1 (middle), and bovine H2B (right). Numbers at the left indicate the H2A.Z sequence position. Hydrophilicity increases to the left (negative values); hydrophobicity to the right (positive values).

DISCUSSION

The histone H2A.Z cDNAs from human, bovine, and rat encode identical proteins. It is evident that the structure of the H2A.Z polypeptide has been highly conserved in species as evolutionarily divergent as sea urchin and man. The H2A.1/2 isoproteins have also been highly conserved during evolution but with only a 59% overall amino acid positional identity to H2A.Z. However, even with the considerable divergence of the H2A.1/2 and H2A.Z isoproteins, several blocks of sequence have been rather highly conserved and many char-

acteristics seem to have been maintained between the two. Combining the data on the locations of allowable changes in H2A.Z sequences with that on the regions of lowest similarity between H2A.Z and H2A.1/2 suggests that the crucial blocks of sequence functionally differentiating these two isoproteins are primarily in two regions of the sequence, H2A.Z positions 73-82 and 97-120. These sequences differ considerably between the two isoproteins, and yet each has been rather highly conserved during evolution. These findings suggest that H2A.Z does function as an H2A in the basic nucleosomal organization (8), but that certain as yet undetermined functional properties differentiate it from H2A.1. H2A.Z cDNAs should be useful tools for investigating the extent that H2A.1/2 and H2A.Z isoproteins can functionally substitute for each other in vivo as well as the nature of the differentiated functions.

A major component in the stability of mRNAs for the replication-linked histone isoproteins is inherent in the message structure (31-36). In general, mRNAs for replication-linked histone isoproteins contain only short 5' and 3' untranslated regions, are not polyadenylated, and contain a conserved stem-loop structure at the 3' end of the mRNA (37,38). The stability of mRNAs in studies with recombinant histone genes and cDNAs has been found to be influenced by the location of the stem-loop structure with respect to the 3' end of the coding sequence and the presence of other gene sequences as well as polyA tailing at the 3' end of the message (31,32,35,39). Ross and coworkers have found that replication-linked histone H4 mRNA is degraded exonucleolytically from its 3' end in cells that have been treated with inhibitors of DNA synthesis (40,41).

In contrast, the mRNAs of the basal histone isoproteins, including those of chicken H2A.F, appear to have longer 3'UTRs with a terminal polyA tail and do not contain the conserved stem-loop structure at the 3' end of the mRNA (11, 12,27,28). H2A.Z cDNAs from all three mammalian species studied also contain lengthy and highly conserved 3'UTRs with an AATAAA polyadenylation signal at the same position. In fact, the mammalian H2A.Z 3'UTRs are more highly conserved than are their coding sequences, suggesting that these sequences have some sort of functional role. In contrast the 3'UTR of the chicken H2A.F (11) has no significant sequence similarity with the mammalian H2A.Z 3'UTRs. Curiously, the situation is quite different for histone H3.3, another histone isoprotein whose synthesis is not linked to DNA replication; the 3'UTRs in these cDNAs are conserved with 85% similarity between mammals and chicken (27,28). The human and rat sequences do have a polyA tail starting 16 nucleotides downstream from AATAAA, but the bovine sequence continues with at least

another 140 nucleotides without a polyA tail. Within the 3'UTR of all three mammalian cDNAs, there is a sequence which consists of a short potential stem structure between two direct repeats of 78% sequence similarity. While no specific function can at present be assigned to this particular sequence, various conserved sequences in the 3'UTRs between the termination codon and the polyadenylation signal of several mRNAs have been reported as possibly being important in the stability and metabolism of those mRNAs (31-33,35-39,42-44). The mammalian H2A.Z cDNA 3'UTR is 64% AT and contains sequences which have been shown to be the kind of sequences that bind various non-histone proteins (43,44).

Although the synthesis of H2A.Z is not linked to DNA replication, it is still regulated with a 20 fold difference in expression between cycling and quiescent cells (3). In this regard, its regulation is perhaps more similar to that of proteins such as the myc oncogene product (45). We are at present analyzing the genomic organization of the gene for H2A.Z as well as the regulation of its expression during different states of cell growth and proliferation.

ACKNOWLEDGEMENTS

Sincere thanks to Sandra Colombini-Hatch (Laboratory of Tumor Cell Biology, NCI, NIH, Bethesda, MD) for providing the human cDNA library, to Jurgen Brosius (Department of Genetics and Development, Columbia University, NY, NY) for providing the bovine and rat cDNA libraries, to Julian Wells (Department of Biochemistry, University of South Australia, Adelaide, South Australia) for providing the chicken H2A.F cDNA and to Amanda Goddard (Laboratory of Medicinal Chemistry, NCI, NIH, Bethesda, MD) for synthesizing most of the oligonucleotides used for sequencing. Concepcion Munses provided excellent technical assistance in carrying out *in vitro* transcription-translation reactions and electrophoresis of the polypeptides in two dimensional gels. Dr. Kurt Kohn of this Laboratory and Dr. Richard Kahn (Laboratory of Biological Chemistry, NCI, NIH) provided assistance in searching through sequence databases for similarities to sequences already on record. Thanks to Madie Tyler for typing the manuscript.

*To whom correspondence should be addressed

REFERENCES

1. West, M.H.P. and Bonner, W.M. (1980) *Biochemistry* **19**, 3238-3245.
2. Wu, R.S. and Bonner, W.M. (1981) *Cell* **27**, 321-330.
3. Wu, R.S., Tsai, S., and Bonner, W.M. (1982) *Cell* **31**, 367-374.
4. Gabrielli, F., Hancock, R., and Faber, A.J. (1981) *Eur. J. Biochem.* **120**, 363-369.
5. 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.
6. Franklin, S.G. and Zweidler, A. (1977) *Nature (London)* **271**, 273-275.
7. West, M.H.P. and Bonner, W.M. (1983) *Comp. Biochem. Physiol.* **76B**, 455-464.

8. Hatch, C.L., Bonner, W.M., and Moudrianakis, E.N. (1983) *Science* **221**, 468-470.
9. Wu, R.S., Panusz, H.T., Hatch, C.L., and Bonner, W.M. (1986) *Crit. Revs. in Biochem.* **20**, 201-263.
10. Wu, R.S., Nishioka, D., and Bonner, W.M. (1982) *J. Cell. Biol.* **93**, 426-431.
11. 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.
12. Ernst, S.G., Miller, H., Brenner, C.A., Nocente-McGrath, C., Francis, S., and McIsaac, R. (1987) *Nucleic Acids Res.* **15**, 4629-4644.
13. Ball, D., Slaughter, C.A., Hensley, P., and Garrard, W.T. (1983) *FEBS Lett.* **154**, 166-170.
14. Nickel, B.E., Roth, S.Y., Cook, R.G., Allis, C.D., and Davie, J.R. (1987) *Biochemistry* **26**, 4427-4421.
15. Colombini, S., Arya, S., Reitz, M., Jagodzinski, L., Beaver, B., Gallo, R., and Wong-Staal, F. (1987) submitted for publication.
16. Okayama, H. and Berg, P. (1982) *Molec. Cell. Biol.* **2**, 161-170.
17. Okayama, H. and Berg, P. (1983) *Molec. Cell. Biol.* **3**, 280-289.
18. Mocchetti, I., Einstein, R., and Brosius, J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7221-7225.
19. Sherbany, A.A., Parent, A.S., and Brosius, J. (1987) *DNA* **6**, 267-272.
20. Kirchhausen, T., Scarmato, P., Harrison, S.C., Monroe, J.J., Chow, E.P., Mattaliano, R.J., Ramachandran, K.L., Smart, J.E., Ahn, A.H., and Brosius, J. (1987) *Science* **236**, 320-324.
21. Feinberg, A. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6-13.
22. Feinberg, A. and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266-267.
23. Henikoff, S. (1984) *Gene* **28**, 351-359.
24. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
25. Bonner, W.M., West, M.H.P., and Stedman, J.D. (1980) *Eur. J. Biochem.* **109**, 17-23.
26. Wickens, M. and Stephenson, P. (1984) *Science* **226**, 1045-1041.
27. Wells, D., Hoffman, D., and Kedes, L. (1987) *Nucleic Acids Res.* **15**, 2871-2889.
28. Dodgson, J.B., Yamamoto, M., and Engel, J.D. (1987) *Nucleic Acids Res.* **15**, 6294.
29. Moss, T., Cary, P.D., Aberchrombie, B.D., Crane-Robinson, C., and Bradbury, E.M. (1976) *Eur. J. Biochem.* **71**, 337-350.
30. Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* **157**, 105-132.
31. Luscher, B., Stauber, C., Schindler, R., and Schumperli, D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4389-4393.
32. Stauber, C., Luscher, B., Eckner, R., Lotscher, E., and Schumperli, D. (1986) *EMBO J.* **5**, 3297-3303.
33. Alterman, R.B., Sprecher, R.A., Graves, R.A., Marzluff, W.F., and Skoultchi, A.I. (1985) *Molec. Cell. Biol.* **5**, 2316-2324.
34. Artishevsky, A., Gafsky, A., and Lee, A.S. (1985) *Science* **230**, 1061-1063.
35. Capasso, O., Bleecker, G.C., and Heintz, N. (1987) *EMBO J.* **6**, 1825-1831.
36. Graves, R.A., Pandey, N.B., Chodchoy, N., and Marzluff, W.F. (1987) *Cell* **48**, 615-626.
37. Birchmeier, C., Folk, W., and Birnstiel, M.C. (1983) *Cell* **35**, 433-440.
38. Birnstiel, M.L., Busslinger, M., and Strub, K. (1985) *Cell* **41**, 349-359.
39. Seiler-Tuyns, A. and Paterson, B.M. (1986) *Nucleic Acids Research* **14**, 8845-8862.
40. Ross, J. and Kobs, G. (1986) *J. Mol. Biol.* **188**, 579-593.
41. Ross, J., Peltz, S.W., Kobs, G., and Brewer, G. (1986) *Molec. Cell. Biol.* **6**, 4362-4371.
42. Shaw, G. and Kamen, R. (1986) *Cell* **46**, 659-667.
43. Reeves, R., Elton, T.S., Nissen, M.S., Lehn, D., and Johnson, K.R. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6531-6535.
44. Solomon, M.J., Strauss, F., and Varshavsky, A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1276-1280.
45. Kelly, K., Cochran, B.H., Stiles, C.D., and Leder, P. (1983) *Cell* **35**, 603-610.