Nucleic Acids Research, 1998, Vol. 26, No. 12 **2837–2842**

Evolutionary conservation of histone macroH2A subtypes and domains

John R. Pehrson* and Reina N. Fuji

University of Pennsylvania, Department of Animal Biology, School of Veterinary Medicine, Philadelphia, PA 19104-6048, USA

Received April 15, 1998; Accepted April 29, 1998

DDBJ/EMBL/GenBank accession nos AF058444–AF058446

ABSTRACT

Histone macroH2A is an unusual core histone that contains a large non-histone region, and a region that resembles a full length H2A. We examined the conservation of this novel structural arrangement by cloning chicken macroH2A cDNAs and comparing them to their rat counterparts. The amino acid sequences of the two known macroH2A subtypes are >95% identical between these species despite evolutionary separation of ~300 million years. The H2A region of macroH2A is completely conserved, and thus is even more conserved than conventional H2A in these species. The origin of the non-histone domain was examined by comparing its sequence to proteins found in bacteria and RNA viruses. These comparisons indicate that this domain is derived from a gene that originated prior to the appearance of eukaryotes, and suggest that the non-histone region has retained the basic function of its ancestral gene.

INTRODUCTION

Core histones are among the most evolutionarily conserved proteins in eukaryotes. This conservation is presumably the result of the critical role that nucleosomes play in DNA packaging and gene regulation. We discovered a new type of core histone, macroH2A (mH2A), in rat liver nucleosomes (1). The N-terminal third of mH2A is 64% identical to a full length H2A. MacroH2A also contains a large region that does not resemble any other known histone (Fig. 1). This large non-histone region distinguishes mH2A from all other known core histones.

Sequences from mH2A cDNAs and reactions with mH2A-specific antibodies established the existence of two distinct mH2A proteins in mammalian tissues (1,2). These subtypes are called mH2A1.1 and mH2A1.2, and they differ from one another in only one region (Fig. 1). The nucleotide sequences of the cDNAs that encode these subtypes are identical in both their coding and non-coding regions except for one short segment in the non-histone region (1,2). This indicates that these subtypes are produced from the same gene by alternate splicing; this has been confirmed by cloning the rat mH2A1 gene (unpublished results). Subtype specific functions are suggested by studies that showed that mH2A1.1 and 1.2 proteins have distinct patterns of expression during development and in different adult organs (2). In rat liver, we estimated that there is one mH2A for every 30 nucleosomes (1). We recently identified a third mH2A subtype that is produced from a separate gene that we call mH2A2 (unpublished results).

The unusual structure of mH2A suggests that it is functionally distinct from conventional H2As. Consistent with this possibility, we recently showed that mH2A is preferentially concentrated in the inactive X chromosome of female mammals (3). This association suggests that mH2A participates in the transcriptional silencing of this chromosome. In the present work we sought to identify the regions of mH2A that are most directly involved in its function(s) by identifying regions that are highly conserved in evolution. We cloned and sequenced chicken mH2A cDNAs, and compared them to those previously known from the rat. These two species separated in evolution ~300 million years ago (4) and prior to the appearance of X chromosome inactivation, which occurred only in mammals (5). We also examined the origins of the H2A and non-histone regions by comparing their sequences to known proteins.

MATERIALS AND METHODS

Cloning and sequencing of chicken macroH2A cDNAs

A chicken liver cDNA library (6) was screened with the non-H2A region of a rat mH2A1.1 cDNA. A positive plaque was identified, and the insert was cloned into pBluescript KS+ (Stratagene). Nested deletions of the mH2A insert were generated using exonuclease III digestion (7). Subclones were sequenced (8) using the Sequenase DNA sequencing kit (US Biochemicals). Both strands were sequenced except for two small segments of the 3′ non-coding region. Separate sequences were generated incorporating either dGTP or dITP as a substitute for dGTP. Reactions with dITP were treated with terminal deoxynucleotidyl transferase to reduce artefacts associated with the use of dITP (US Biochemicals). This chicken mH2A cDNA was missing the region of non-identity between mH2A1.1 and 1.2, and therefore encodes a truncated mH2A protein. Attempts to confirm the expression of this truncated mH2A gave ambiguous results.

The polymerase chain reaction (PCR) was used to amplify mH2A cDNAs that contain the region that was missing from the

*To whom correspondence should be addressed. Tel: +1 215 898 0454; Fax: +1 215 898 9923; Email: pehrson@vet.upenn.edu

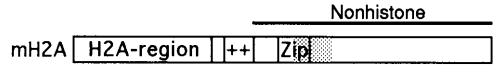

Figure 1. Diagram of the structure of mH2A subtypes. ++ indicates a lysine-rich region that resembles part of the C-terminal domain of histone H1, Zip indicates a region that resembles a leucine zipper, and the gray region shows the location of the region that is different between mH2A1.1 and 1.2 (1,2). The region C-terminal to the lysine-rich region is referred to as the non-histone region (residues 160–367 of rat mH2A1.1). Accession numbers for rat mH2A1.1 and 1.2 are M99065 and U79139, respectively.

cDNA clone discussed above. One ng of cDNA from chicken liver, brain, spleen or muscle served as a template and amplification was achieved by 30 cycles of 1 min at 95°C, 30 s at 50°C and 2 min at 72°C. The reactions were carried out in a standard reaction buffer (Promega) containing 1.5 mM $MgCl_2$, and Tli DNA polymerase (Promega) was used to minimize mutations during amplification. The primer sequences were GGAATTCCAA-GAAGCAGGGAGAAGT and GGAATTCCACAAACTCCTTG-CCGCC; these sequences include sites for restriction nuclease cleavage used in cloning the products. A small amount of the PCR products was radiolabeled using T4 polynucleotide kinase and run on a 6% denaturing acrylamide gel for analysis (7). The remaining DNA was cloned into pBluescript KS+ and sequenced. Both strands of two independent clones were sequenced for each reported sequence. The major products of these PCRs contained the region that was missing from the original chicken cDNA. Many of the PCRs also produced a minor product of the size expected for cDNAs that lack this region (see liver sample in Fig. 2); the predicted size for such a product is 164 bp. Attempts to reamplify and clone this fragment were unsuccessful.

Antibody production and purification

The non-histone region of rat mH2A1.1 (residues 160–367) was expressed in *Escherichia coli* (strain BL-21) as a glutathione transferase fusion protein using the expression vector pGEX-2TK (9,10). The fusion protein was purified using glutathione agarose beads (10). Antibodies against the fusion protein were raised in chickens. IgY was prepared from egg yolks (11) and immuno-affinity purified (2).

Western blot analysis

Frozen chicken liver was obtained from Pel-Freez Biologicals. Adult and embryonic chicken blood were from Hy-Vac laboratories. The nuclei were isolated and digested with micrococcal nuclease (12). An equal volume of 2× SDS sample buffer was added to the digests, they were run in SDS polyacrylamide gels and western blots were performed (2).

Phylogenetic analysis of H2A protein sequences

Histone H2A protein sequences were obtained from the histone sequence database (13) (http://www.nhgri.nih.gov/DIR/GTB/HISTONES/). The 71 H2A sequences previously used to construct a phylogenetic tree of H2As (14) were aligned along with the H2A region of mH2A using Clustal W (15) (http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html). The aligned sequences were analyzed using Joseph Felsenstein's phylogeny inference package PHYLIP (http://bioweb.pasteur.fr/seqanal/interfaces/phylip-uk.html); distance measures were calculated from the aligned protein sequences with PROTDIST using maximum likelihood estimates based on the Dayhoff PAM matrix, and the phylogenetic tree was produced from the distances matrix by NEIGHBOR using Saitou and Nei's 'Neighbor Joining Method'. The phylogenetic tree was viewed with TreeView 1.5 (16) (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

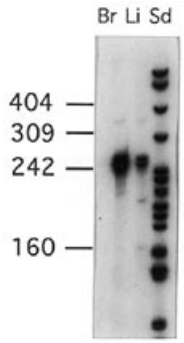

Figure 2. PCR amplification of the regions of non-identity of chicken mH2A1.1 and 1.2. Primers that flank the region that is different between mH2A1.1 and 1.2 were used in PCRs that used cDNA from chicken brain (lane Br) or liver (lane Li) as a template. Reaction products were labeled with ^{32}P and run on a denaturing acrylamide gel. Lane Sd, end-labeled *Msp*I digested pBR322 DNA. Numbers on the left indicate the length of selected marker fragments.

RESULTS

Identification and sequencing of chicken mH2A cDNAs

A chicken liver cDNA library was screened with a cDNA fragment from the non-histone region of rat mH2A1.1. A positive clone contained a 1720 bp insert that was completely sequenced and found to be highly homologous to rat mH2A. However, this chicken cDNA was missing the region that is different between the two known rat mH2A subtypes, mH2A1.1 and 1.2 (gray region in Fig. 1). To obtain these missing sequences, segments of chicken cDNAs that contain them were amplified by PCR. Two prominent products were made when cDNA from chicken liver, brain, spleen or muscle was used as a template for the PCR; the results with liver and brain cDNA are shown in Figure 2. Sequencing of these products revealed that they contain the missing regions of mH2A1.1 and 1.2. Surprisingly, the nucleotide sequences of these regions are nearly identical to those of the rat: 96% for mH2A1.1 and 98% for mH2A1.2 (Fig. 3). The nucleotide sequences of the PCR products outside this conserved region are identical to the original chicken cDNA and differ from the homologous rat sequences at 25 of 114 positions (parts of these sequences are shown in Fig. 2). This demonstrates that these PCR products were made from chicken cDNA, and not trace contaminants of rat cDNAs. The exceptional conservation of the nucleotide sequences in the region that is different between mH2A1.1 and 1.2 does not occur in other regions. Overall, the nucleotide sequences of the coding regions of rat and chicken mH2A1.1 are 83% identical. The non-coding regions are highly divergent.

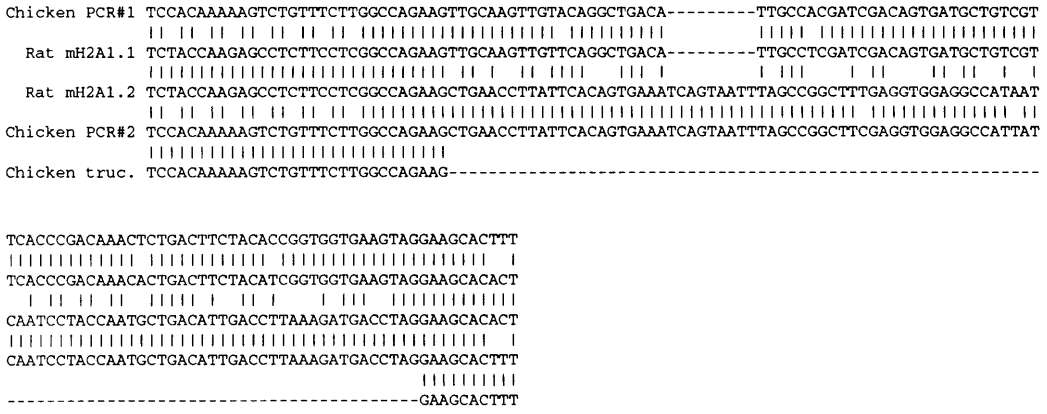

Figure 3. Nucleotide sequences of the regions that are different between mH2A1.1 and 1.2. Sequences of the products of the PCRs shown in Figure 2 are compared to the regions that are different between rat mH2A1.1 and 1.2. The complete sequences of chicken PCR#1 (accession no. AF058445) and chicken PCR#2 (accession no. AF058446) are in GenBank. The sequence labeled Chicken trunc. is from the chicken liver cDNA that is missing this region (accession no. AF058444). (|) indicates identities and (-) indicates a gap.

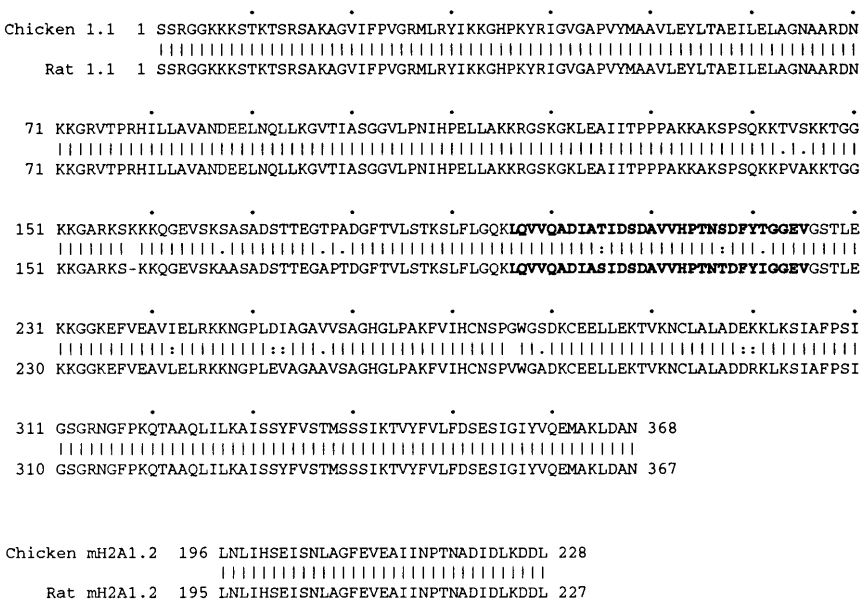

Figure 4. Comparison of rat and chicken mH2A subtypes. Amino acids are indicated by their one letter code. The sequences for mH2A1.1 are shown and the region of non-identity is in bold. The region of non-identity of mH2A1.2 is shown below. (|) indicates identities, (:) indicates conservative substitutions, (.) indicates less conservative substitutions and (-) indicates a gap.

Comparison of chicken and rat mH2A subtypes

The amino acid sequences of chicken mH2A1.1 and 1.2 were deduced from the cDNA sequences and are compared to rat mH2As in Figure 4. Both subtypes have been remarkably conserved over their entire length. Overall, the conservation is 95% for mH2A1.1 and 96% for mH2A1.2. The H2A region (residues 1–122) is completely conserved, and thus is even more conserved than conventional H2A in these species; conventional chicken and mammalian H2As differ from one another at two to four positions depending on which subtypes are compared (17). The complete conservation of the H2A region of mH2A is consistent with its evident role in nucleosome core formation (1). The non-histone region of mH2A (residues 160–367 in rat mH2A1.1; Fig. 1) is also exceptionally conserved: 93% identical for mH2A1.1 and 95% for mH2A1.2. This degree of conservation approaches that of the core histones H2A and H2B [98 and 97% identical, respectively, between chickens and rats (17)], and is significantly greater than most proteins. This suggests that most of the non-histone region is involved in specific interactions with nuclear components or is unusually constrained by its three-dimensional structure.

Expression of chicken mH2A

Western blots were used to examine the expression of mH2A proteins in chicken liver and blood. Antibodies against the non-histone region of rat mH2A1.1 detected two proteins in

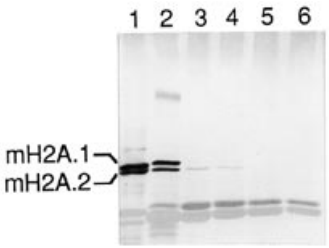

Figure 5. Western blot detection of mH2A in chicken liver and blood. Proteins were extracted from nuclei and stained using an affinity purified antibody specific for the non-histone region of rat mH2A1.1. Nuclei were prepared from: 1, rat liver; 2, chicken liver; 3, 5-day embryonic chicken blood; 4, 8-day embryonic chicken blood; 5, 14-day embryonic chicken blood; and 6, adult chicken blood. The bands below mH2A are histone H1, which crossreacted with this antiserum for unknown reasons. Loadings were adjusted to equalize amount of core histone present in each lane.

chicken liver nuclear extracts with electrophoretic mobilities virtually identical to rat mH2A1.1 and 1.2 (Fig. 5, lane 2). The relative intensity of the mH2A bands in chicken liver was lower than in rat liver. However, the level of chicken mH2A may be underestimated since this antisera was raised against rat mH2A. In addition, chicken liver contains nucleated red blood cells which contribute chromatin, but have little or no mH2A (see below).

Neither mH2A1.1 nor 1.2 were detected in nuclear extracts from adult chicken blood (Fig. 5, lane 6). Nuclei from blood of 5- and 8-day embryos contained a low level of mH2A1.2, but no mH2A1.1 (Fig. 4, lanes 3 and 4); at these stages of development, immature red cells are abundant in blood (18). In 14-day embryos most red blood cells have reached developmental maturity (18) and neither mH2A subtype was detected in nuclei isolated from blood at this stage (Fig. 5, lane 5). These results indicate that mH2A is expressed in the erythrocyte lineage, but is absent or present at very low levels in mature erythrocytes. Using conventional core histones as an internal standard, we estimated that the mH2A content of mature chicken red blood cells is at least 20-fold lower than chicken liver; this estimate is based on western blots of serial dilutions of the chicken liver nuclear extracts.

Comparisons of the non-histone and H2A regions of mH2A to known proteins

The non-histone region of mH2A is homologous to a protein encoded by a gene found in some bacteria. One of these genes was discovered serendipitously in the bacteria *Alcaligenes eutrophus* (19), and we identified a homologous gene in the genomic database of *E.coli*. Alignment of the non-histone region of mH2A1.1 with these bacterial proteins is shown in Figure 6A. The homology extends across nearly the entire length of the bacterial proteins, though the bacterial proteins lack a region corresponding to the first 30 amino acids of the non-histone region. The non-histone region of mH2A1.1 is 34 and 30% identical to the *A.eutrophus* and *E.coli* proteins, respectively.

We found that the non-histone region is also homologous to part of a protein of some positive-strand RNA viruses. Interestingly, the region of homology corresponds to a segment of ~100 amino acids which is the most conserved region between the alphaviruses and rubella virus (20,21). An alignment of this region of the alphavirus sindbis virus and rubella virus to the corresponding region of mH2A1.1 is shown in Figure 6B. The non-histone region

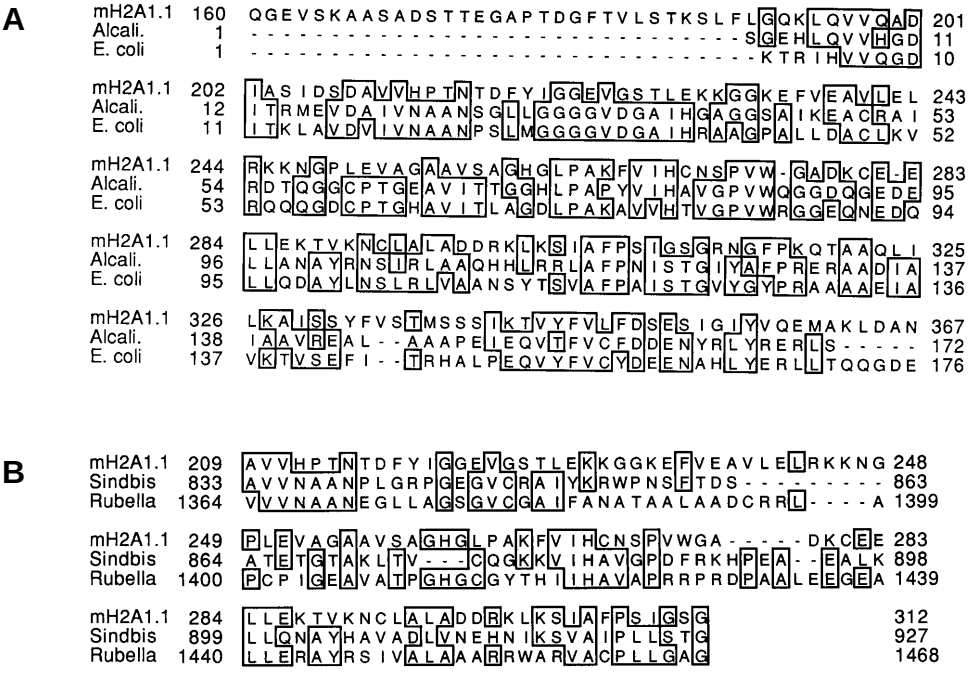

Figure 6. Comparison of the non-histone region of mH2A to bacterial and viral proteins. (**A**) Alignment of the complete non-histone region of rat mH2A1.1, residues 160–367, to proteins from *A.eutrophus* (accession no. L36817), and *E.coli* (accession no. 1787283). (**B**) Alignment of residues 209–312 of the non-histone region of mH2A1.1 with residues 1364–1459 of the non-structural polyprotein of Sindbis virus (P03317 Swiss-Prot) and residues 833–938 of the non-structural polyprotein of Rubella virus (P13889 Swiss-Prot). Identical residues are boxed, and (-) indicates a gap.

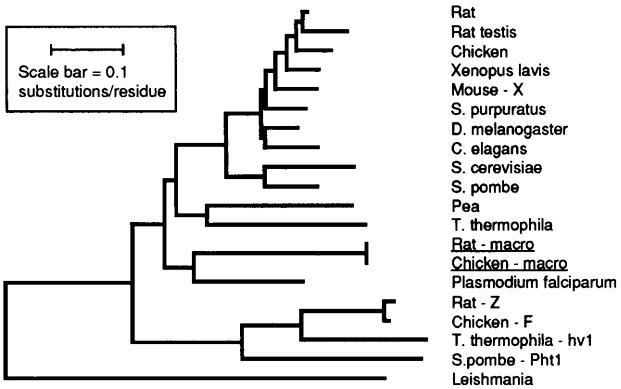

Figure 7. Phylogenetic tree analysis of the H2A region of mH2A with other H2As. The H2A region of mH2A was aligned with 71 H2A sequences previously used in a phylogenetic analysis (14). A phylogenetic tree was constructed from these aligned sequences (Materials and Methods). Only selected representatives of the major groups are shown. The complete tree was essentially identical to one previously published (14). The accession numbers of the H2As shown are: Rat, A02591; Rat testis, X59962; Chicken, V00413; *Xenopus*, M21287; Mouse H2A.X, X58069; *Strongylocentrotus purpuratus*, X06642; *Drosophila melanogaster*, S10094; *Caenorhabditis elegans*, X15633; *S.cerevisiae*, V01304; *Schizosaccharomyces pombe*, X05220; Pea, JQ1183; *T.thermophila*, L18892; Rat macroH2A, U79139; Chicken macroH2A, AF058444; *P.falciparum*, M86865; Rat H2A.Z, M37584; Chicken H2A.F, V00414; *T.thermophila* H2A.hv1, X15548; *S.pombe* H2A.Pht1, S52560; *Leishmania*, X60054.

is 24 and 25% identical to this domain of sindbis and rubella viruses, respectively. These two viral sequences are 36% identical to one another and are ~40% identical to the bacterial proteins (Table 1). Thus, all of these proteins appear to be related to one another.

The H2A region of mH2A almost certainly arose from an H2A gene. We examined the relationship of the H2A region to known H2As by constructing a phylogenetic tree (Fig. 7). This analysis did not reveal a close link between the H2A region and any conventional H2A or H2A variant. The nearest link is to the H2A of the malaria parasite *Plasmodium falciparum.*

DISCUSSION

The remarkable conservation of both mH2A1.1 and 1.2 between chickens and rats indicates that the basic function(s) of these mH2A subtypes have been conserved during the 300 million years of evolution that separate birds and mammals. Our recent studies showing that mH2A is preferentially concentrated in the inactive X chromosome of female mammals suggest that mH2A is involved in transcriptional silencing (3). Since X chromosome inactivation arose in the mammalian lineage after mammals separated from birds (5), the specific role of mH2A in X-inactivation is most likely an adaptation of a pre-existing mH2A function. One interesting possibility is that mH2A participates in gene silencing of autosomal regions in both birds and mammals, and that this function was adapted to X-inactivation in mammals. This possibility is consistent with the finding that mH2A is present in autosomes of both male and female mammals (3). The observation that mH2A is absent or at very low levels in mature chicken erythrocytes indicates that it is not involved in the transcriptional silencing that occurs in these cells.

Table 1. Percentage identity of non-histone region of mH2A1.1, and bacterial and viral homologues

	mH2A	*Alcal.*	*E.coli*	Sindbis
Alcaligenes	34			
E.coli	30	55		
Sindbis	24	40	40	
Rubella	25	41	35	36

Regions of comparison are those shown in Figure 6.

The complete conservation of the H2A region of mH2A between birds and mammals is interesting since this region is only 64% identical to a conventional rat H2A (1). This suggests that the differences between conventional H2A and the H2A region of mH2A are functionally significant, a possibility consistent with the observation that some core variants are functionally distinct from their conventional counterparts. One example is H2A.Z, an H2A that is 59% identical to conventional H2A (22). Recent studies showing that H2A.Z is essential in *Drosophila* (23) and *Tetrahymena thermophila* (24) indicate that it has important function(s) that cannot be carried out by conventional H2A. Another example is CENP-A, a 17 kDa centromere-specific protein (25) that co-purifies with mononucleosomes (26) and has a 93 amino acid domain that is 62% identical to histone H3 (27,28). This H3-like domain can localize to centromeres (28), showing that a variant core histone domain can be targeted to a specific chromosomal region.

The non-histone region of mH2A appears to have originated from a gene that existed prior to separation of eubacteria and eukaryotes. This is indicated by the existence of eubacterial proteins that are homologous to the non-histone region. The degree of homology (Table 1) is similar to the average of 37% identity observed for 57 enzymes conserved between eukaryotes and eubacteria (29), suggesting that the basic function of the non-histone region and these prokaryotic homologues is very similar. Unfortunately, the function of the bacterial homologues is not known.

A potential clue to the function of the non-histone region comes from its homology to a domain found in RNA viruses (Fig. 6). Although the function of this viral domain is not known, it is the most conserved sequence between sindbis virus and rubella virus (20,21). In sindbis virus the domain is part of a protein that associates with viral RNA and is required for the synthesis of negative-strand RNA (30,31). One interesting possibility in terms of mH2A function is that this domain binds RNA; it seems less likely that this viral domain binds DNA since these viruses have no DNA intermediates. There is mounting evidence that nuclear RNAs can participate in regulating chromatin structure and function, possibly through a direct interaction between RNA and chromatin (32). The best described example of this is X chromosome inactivation, which involves a nuclear RNA from the *Xist* gene. The preferential association of mH2A with the inactive X chromosome could involve an interaction of the non-histone region with *Xist* RNA.

The mH2A gene appears to have formed by the linking of an H2A gene to a non-histone gene. The divergence of the H2A region from conventional H2As suggests that the mH2A gene was formed relatively early in eukaryotic evolution. A phylogenetic analysis of the H2A region of mH2A with other known H2As

suggests that the H2A region branched from the H2A phylogenetic tree just prior to the branching of plants and animals. Assuming that this separation corresponds to the formation of the primordial mH2A gene, this analysis suggests that mH2A could potentially be present in many eukaryotes including yeast. However, a search of the complete genome of the yeast *Saccharomyces cerevisiae* failed to find any sequence that resembles mH2A. It is possible that mH2A was lost in the evolution of this yeast. Alternatively, the phylogenetic analysis could be misleading if the H2A region of mH2A had a period where it evolved more rapidly than conventional H2As. Such a period could have occurred shortly after the formation of mH2A. At that point the H2A region would probably not have been constrained like a conventional H2A, and would not have acquired specialized structures related to mH2A function. In this scenario the mH2A gene may have appeared more recently than indicated by the phylogenetic analysis.

MacroH2A appears to be the only known core histone that contains a large domain derived from a non-histone gene. The H3 variant CENP-A contains a highly divergent 47 amino acid N-terminal domain. However, this domain is the same size as the corresponding N-terminal region of H3, and like the conventional N-terminus of H3, is rich in basic amino acids (28). Thus, it seems likely that this domain evolved from the N-terminus of a conventional H3. Although there may be other examples of core histones that became linked to a non-histone gene, clearly they are rare. This is not surprising given the extreme constraints imposed on the structures of the core histones. The high conservation of mH2A structure seen in this study indicates that this combination of core histone and non-histone domains has acquired valuable functions.

ACKNOWLEDGEMENTS

We thank Colin MacNeill and John Burch for providing chicken cDNAs and assisting with the cDNA library, Chhaya Dharia for preparation and purification of antibodies, Carl Costanzi for the phylogenetic analysis and reading the manuscript, Leslie Taylor for technical assistance, William Kaelin, Jr for providing the plasmid pGEX-27 and Mike Hall for providing embryonic chicken blood. This work was supported by NIH grant GM49351 and a grant from Merck.

REFERENCES

1 Pehrson, J. R. and Fried, V. A. (1992) *Science*, **257**, 1398–1400.
2 Pehrson, J. R., Costanzi, C. and Dharia, C. (1997) *J. Cell. Biochem.*, **65**, 107–113.
3 Costanzi, C. and Pehrson, J. R. (1998) *Nature*, in press.
4 McLaughlin, P. J. and Dayhoff, M. O. (1969) *Atlas Protein Seq. Struct.*, **4**, 39–46.
5 Solari, A. J. (1994) *Sex Chromosomes and Sex Determination in Vertebrates.* CRC Press, Boca Ratan.
6 Kirchgessner, T. G., Heinzmann, C., Scenson, K. L., Gordon, D. A., Nicosia, M., Lebherz, H. G., Lusis, A. J. and Williams, D. L. (1987) *Gene*, **59**, 241–251.
7 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, (2nd ed.). Cold Spring Harbor Laboratory Press, Plainview, NY.
8 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
9 Kaelin, W. G. J., Pallas, D. C., DeCaprio, J. A., Kaye, F. J. and Livingston, D. M. (1991) *Cell*, **64**, 521–532.
10 Smith, D. B. and Johnson, K. S. (1988) *Gene*, **67**, 31–40.
11 Gassmann, M., Thömmes, P., Weiser, T. and Hübscher, U. (1990) *FASEB J.*, **4**, 2528–2532.
12 Pehrson, J. R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9149–9153.
13 Baxevanis, A. D. and Landsman, D. (1997) *Nucleic Acids Res.*, **25**, 272–272.
14 Thatcher, T. H. and Gorovsky, M. A. (1994) *Nucleic Acids Res.*, **22**, 174–179.
15 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) *Nucleic Acids Res.*, **22**, 4673–4680.
16 Page, R. D. M. (1996) *Comput. Appl. Biosci.*, **12**, 357–358.
17 Wells, D. and McBride, C. (1989) *Nucleic Acids Res.*, **17**, Suppl, 311–346.
18 Bruns, G. A. and Ingram, V. M. (1973) *Phil. Trans. R. Soc. Lond., Ser. B.: Biological Sciences*, **266**, 225–305.
19 Valentin, H. E., Zwingmann, G., Schonebaum, A. and Steinbuchel, A. (1995) *Eur. J. Biochem.*, **227**, 43–60.
20 Dominguez, G., Wang, C.-Y. and Frey, T. K. (1990) *Virology*, **177**, 225–238.
21 Gorbalenya, A. E., Koonin, E. V. and Lai, M. M.-C. (1991) *FEBS Lett.*, **288**, 201–205.
22 Hatch, C. L. and Bonner, W. M. (1988) *Nucleic Acids Res.*, **16**, 1113–1124.
23 van Daal, A. and Elgin, S. C. R. (1992) *Mol. Biol. Cell*, **3**, 593–602.
24 Liu, X., Li, B. and Gorovsky, M. A. (1996) *Mol. Cell. Biol.*, **16**, 4305–4311.
25 Earnshaw, W. C. and Rothfield, N. (1985) *Chromosoma*, **91**, 313–321.
26 Palmer, D. K., O'Day, K., Wener, M. H., Andres, B. S. and Margolis, R. L. (1987) *J. Cell Biol.*, **104**, 805–815.
27 Palmer, D. K., O'Day, K., Trong, H. L., Charbonneau, H. and Margolis, R. L. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3734–3738.
28 Sullivan, K. F., Hechenberger, M. and Masri, K. (1994) *J. Cell Biol.*, **127**, 581–592.
29 Doolittle, R. F., Feng, D.-F., Tsang, S., Cho, G. and Little, E. (1996) *Science*, **271**, 470–477.
30 LaStarza, M. W., Lemm, J. A. and Rice, C. M. (1994) *J. Virol.*, **68**, 5781–5791.
31 Wang, Y., Sawicki, S. G. and Sawicki, D. L. (1994) *J. Virol.*, **68**, 6466–6475.
32 Willard, H. F. and Salz, H. K. (1997) *Nature*, **386**, 228–229.