MOLECULAR AND CELLULAR BIOLOGY, Sept. 2005, p. 7616–7624 Vol. 25, No. 17
0270-7306/05/$08.00+0 doi:10.1128/MCB.25.17.7616–7624.2005

Structural Characterization of the Histone Variant macroH2A

Srinivas Chakravarthy,[1] Sampath Kumar Y. Gundimella,[1] Cecile Caron,[2] Pierre-Yves Perche,[2] John R. Pehrson,[3] Saadi Khochbin,[2] and Karolin Luger[1]*

Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523-1870[1]; INSERM U309-Laboratoire de Biologie Moléculaire et Cellulaire de la Différenciation-Équipe Chromatine et Expression des genes, Institut Albert Bonniot, Faculté de médecine, Domaine de la Merci, 38 706 La Tronche, France[2]; and Department of Animal Biology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104[3]

Received 23 March 2005/Returned for modification 15 April 2005/Accepted 6 June 2005

macroH2A is an H2A variant with a highly unusual structural organization. It has a C-terminal domain connected to the N-terminal histone domain by a linker. Crystallographic and biochemical studies show that changes in the L1 loop in the histone fold region of macroH2A impact the structure and potentially the function of nucleosomes. The 1.6-Å X-ray structure of the nonhistone region reveals an α/β fold which has previously been found in a functionally diverse group of proteins. This region associates with histone deacetylases and affects the acetylation status of nucleosomes containing macroH2A. Thus, the unusual domain structure of macroH2A integrates independent functions that are instrumental in establishing a structurally and functionally unique chromatin domain.

The compaction of DNA into chromatin is an important regulator of DNA accessibility. The nucleosome core particle (NCP), the fundamental repeating unit of chromatin, plays a central role in the regulation of transcription, replication, and repair. An important emerging mechanism to alter the fundamental biochemical composition and characteristics of chromatin is the substitution of major-type core histones with histone variants (18). This may be achieved by structural alterations in the NCP and/or in chromatin higher-order structures that are brought about by the amino acid sequence differences between the histone variants and their corresponding core counterparts (9; for an example, see reference 28). macroH2A1, with a molecular weight of ~40 kDa, is almost three times the size of major, replication-dependent H2A and is unique among known histone variants due to its unconventional tripartite structural organization (23). The N-terminal third of its amino acid sequence (amino acids [aa] 1 through 122) is 64% identical to major H2A. A C-terminal nonhistone region (aa 161 through 371) is linked to the histone homology domain via a linker region (aa 123 through 160) (Fig. 1A). The C-terminal nonhistone region in itself exhibits amino acid similarities to members of a functionally highly diverse group of proteins that exist in organisms ranging from bacteria and archaea to eukaryotes, and its function remains unknown (24). macroH2A preferentially localizes at the inactive X-chromosome of mammalian female cells, where it may contribute to the maintenance of transcriptionally silent chromatin (7). Recent studies indicate that macroH2A-containing nucleosomes are repressive toward transcription (4, 25). Here, we have combined X-ray crystallography with biochemical and mutational studies to better understand the biological function of macroH2A.

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523-1870. Phone: (970) 491-6405. Fax: (970) 491-0494. E-mail: kluger@lamar.colostate.edu.

MATERIALS AND METHODS

Expression and purification of histone proteins and reconstitution of nucleosomes. All histones were overexpressed in BL21(DE3)-plysS (Stratagene) and purified using previously published protocols (17). The histone domain of macroH2A (aa 1 to 120; macroH2A-HD), together with mouse H2B, H3, and H4, was refolded to a histone octamer (macrooctamer). This was reconstituted onto a 146-bp palindromic DNA fragment derived from human α-satellite regions (α-sat DNA) (16) using salt gradient dialysis (8), resulting in macro-NCP. Milligram amounts of macro-NCP were heat shifted and purified by preparative gel electrophoresis using published protocols (8).

Crystallographic procedures for macro-NCP. macro-NCP was crystallized using salting in vapor diffusion at NCP concentrations ranging from 8 to 12 mg/ml with salt concentrations of 34 to 37.5 mM KCl and 40 to 45 mM $MnCl_2$. The crystals were soaked in 24% 2-methyl-2,4-pentanediol–5% trehalose and frozen in liquid nitrogen as described previously (16). Data were collected at Advanced Light Source (Lawrence Berkeley National Laboratory) on Beamline 8.2.2. Data from a single crystal were processed using Denzo and Scalepack (22). Molecular replacement was performed using Protein Data Bank entry 1AOI as the search model. Molecular replacement and subsequent rounds of refinement were performed using a crystallography and nuclear magnetic resonance system (CNS) (27). The program O was used for model building (11). The veracity of the model was checked using SA-OMIT maps for critical regions during various stages of refinement and a composite omit map at the end.

Expression and purification of the nonhistone region of macroH2A. The coding sequence for amino acids 180 through 367 of macroH2A was subcloned into pGEX4T2 and transformed into BL21(DE3)pLysS. The transformed cells were used to inoculate 6 ml of 2× tryptone-yeast extract medium in the presence of ampicillin (50 μg/ml), chloramphenicol (34 μg/ml), and 5% glucose and grown to turbidity. This primary culture was transferred to a 100-ml culture in the presence of the same drugs as above and 5% glucose, grown for 1.5 to 2 h, and then amplified to 3 liters. The culture was then allowed to grow at 37°C until it reached an optical density (at 600 nm) of 0.4 to 0.6. Expression was induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and the protein was expressed overnight at 25°C. Bacteria were harvested by centrifugation and were resuspended in 1/10 the volume of lysis buffer (150 mM NaCl, 5 mM EDTA, 20 mM Tris HCl [pH 7.5], 1 mM dithiothreitol [DTT], 1 mM benzamidine). This suspension was then flash frozen in liquid nitrogen and stored at −20°C. The cell suspension was thawed at 37°C and sonicated, the insoluble portion was spun down, and the supernatant was loaded on 1/10 the volume of preswollen gluta-

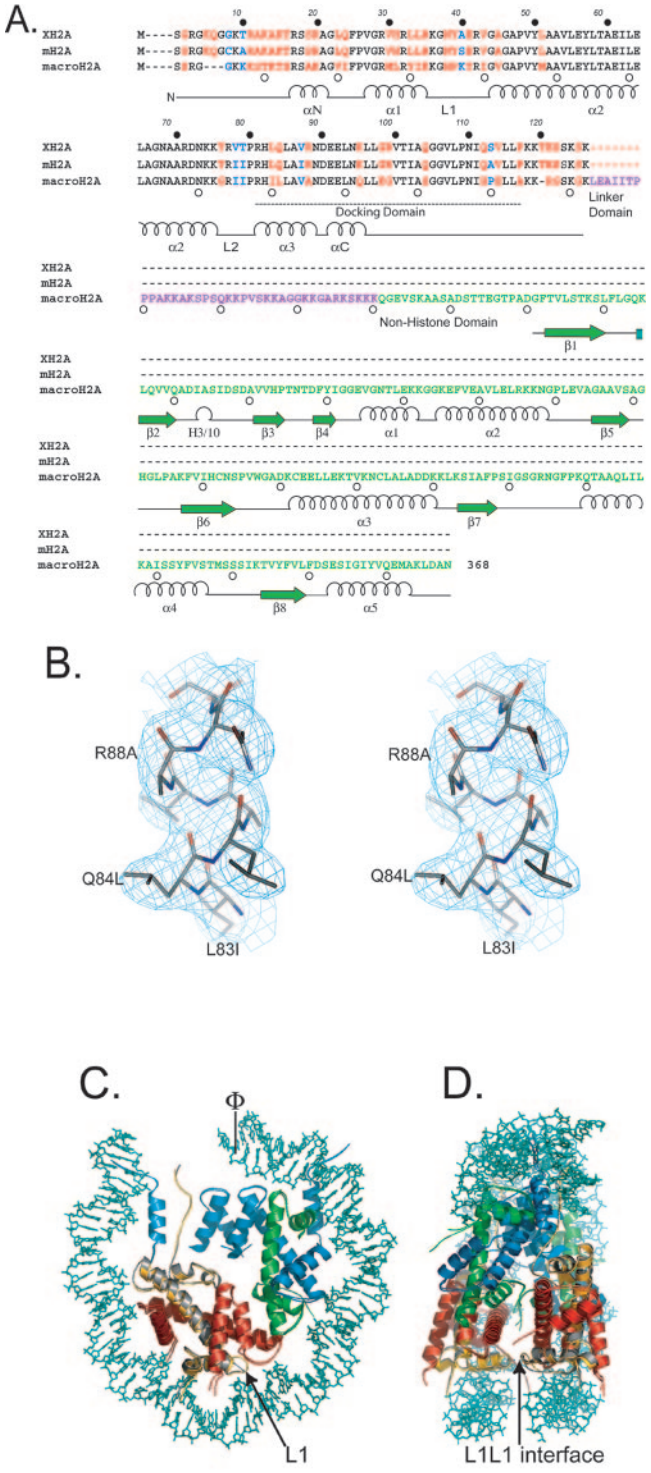

FIG. 1. The overall structure of macro-NCP is similar to that of major NCP. (A) Sequence alignment of *Xenopus laevis* H2A, mouse H2A, and full-length human macroH2A. Filled circles indicate intervals of 10 amino acids in major H2A. Open circles indicate intervals of 10 amino acids in macroH2A. Differences between major H2A (mouse and *X. laevis* or macroH2A and *Xla*-H2A, respectively) and macroH2A are shown in red. Differences between mouse and *Xla*-H2A are shown in blue. The linker region and the nonhistone region of macroH2A are shown in violet and green, respectively. The secondary structure elements of the histone fold (α1, α2, and α3) and extensions (αN and αC) are indicated, as are secondary structure elements for the nonhistone region. The fine broken line delineates the docking domain. (B) Stereo view of a section of the 2Fo-Fc electron density map, calculated at 3Å and contoured at 1σ, showing sequence differences between macroH2A and Xla-H2A L83 to I80, Q84 to L81, and R88 to A85. (C) Superposition of major NCP and macro-NCP (only histone octamers are superimposed) viewed down the superhelical axis. Only 73 bp of the DNA and associated proteins are shown. The central base pair is indicated (ϕ). H3 is colored blue, H4 green, H2B red, H2A yellow, macroH2A gray, and DNA turquoise. The L1 loop is indicated. (D) Side view of the superimposed nucleosomes in panel C rotated 90° around the *y* axis with parts of the DNA removed for clarity. The L1-L1 interface is indicated.

thione *S*-transferase–agarose beads and rocked overnight at 4°C. Beads were washed in 3 volumes of phosphate-buffered saline, followed by 3 volumes of Pre Scission protease buffer (50 mM Tris HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA, 1 mM DTT). Pre Scission protease (60 to 70 μl for every ml of swollen beads) was added, and the beads were rocked overnight at 4°C. The supernatant contains the nonhistone region of macroH2A without the glutathione *S*-transferase tag. The beads were washed with 3 volumes of Pre Scission protease buffer. All the washes and the elute were then analyzed on a sodium dodecyl sulfate–18% polyacrylamide electrophoresis gel. The protein was concentrated to 5 to 10 mg/ml and loaded on a Superdex-75 10/30 (size exclusion) column. Fractions were pooled and concentrated to ~25 mg/ml to be used in crystallization screens.

Crystallographic procedures for the nonhistone region. The nonhistone region of macroH2A (aa 180 to 367) was crystallized with 28% PEG 2000–0.2 M ammonium sulfate–0.1 M sodium acetate, pH 5.9. Heavy atom derivatives were obtained by cocrystallizing the nonhistone region of macroH2A with potassium dicyanoaurate [$KAu(CN)_2$]. Data were collected for both native and gold-derivatized crystals at Advanced Light Source (Lawrence Berkeley National Laboratory) on Beamline 8.3.1 to resolutions of 1.6 Å and 2.1 Å, respectively. Phases were obtained by Multiple-Wavelength Anomalous Dispersion in a two-wavelength experiment. Phases were extended to 1.6 Å using the data from the native crystal. Denzo and Scalepack were used to index and scale the data (22). Experimental phases were obtained using the program SOLVE/RESOLVE (31). We then performed multiple rounds of refinement and manual rebuilding using the CNS (27) and O (11) programs.

Plasmids for immunoprecipitation. Gal4 nonhistone region (G4-NHR) fusion constructs were done as follows: the coding sequence of aa 160 to 371 or 180 to 371 was amplified by PCR with primers creating restrictions sites and inserted in the pCDNA-Gal4 plasmid. Site-directed mutations were introduced with the QuikChange site-directed mutagenesis kit (Stratagene), using G4-NHR 160-371 and G4-NHR 180-371 DNA as templates.

Gene reporter assays. Cos cells were transfected by the calcium phosphate coprecipitation method with reporter plasmids pGal4-TK-luciferase (1 μg) and pCMV-β-gal (200 ng) as internal controls and 0.5 μg of expression vectors for G4-NHD constructs, Gal4 DNA-binding domain alone (G4-BD), or an empty vector. Forty-eight hours after transfection, luciferase and beta-galactosidase activities were measured on Cos extracts using the luciferase (Promega) and beta-galactosidase (Clontech) reporter assay systems according to the manufacturer's instructions. Luciferase values were normalized with respect to beta-galactosidase values.

Immunoprecipitation. Cos cells transfected by the calcium phosphate coprecipitation method with 3 μg of Gal4 fusion constructs and 3 μg of pFlag HDAC1 were lysed 48 h after transfection in LSDB buffer (20% glycerol, 3 mM $MgCl_2$, 50 mM HEPES [pH 7.9], 0,1% NP-40, 1 mM DTT, and protease inhibitor cocktail [Complete Mini EDTA-free medium; Roche] containing 500 mM or 800 mM KCl [LSDB500 or LSDB800, respectively]). After centrifugation, the lysate was incubated with 1 μg of anti-Gal4 antibody (RK5C1; Santa Cruz) for 1 h on ice. Protein G-Sepharose beads (Amersham) were then added and incubated at 4°C for 1 h. After 3 washes with LSBD500 or LSDB800, complexes were recovered by adding Laemmli sample buffer and analyzed by Western blot analysis.

Immunoprecipitation of macroH2A-associated chromatin was performed similarly, except that cells transfected with 5 μg of expression vectors for hemagglutinin (Ha)-H2A or Ha-macroH2A were lysed in radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH 7.4], 1% NP-40, 0.25% Na-deoxycholate, 1 mM DTT, protease inhibitor cocktail) followed by sonication (100 J) in order to obtain chromatin fibers around 600 to 1,000 bp and that immunoprecipitations (IPs) were performed with an anti-Ha antibody (high-affinity 3F10; Roche). Washes of immunoprecipitated complexes were performed in radioimmunoprecipitation assay buffer.

TABLE 1. Data collection and refinement statistics

Statistic (units)	Value or range
Data collection	
Space group	$P2_12_12_1$
Unit cell dimensions	a = 105.5, b = 109.6, c = 176.0
Resolution range (Å)	50–2.95
No. of unique reflections	43,366
% Completeness (overall/highest-resolution shell)	99.5/100
% R_{merge}[a] (overall/last shell)	9.5/42
Refinement	
No. of amino acid residues in the final model[b]	759
No. of base pairs in the DNA	146
No. of water molecules	105
Total no. of atoms in the final model	11,952
R-factor[c]/R_{free}	0.206/0.260
Resolution range (Å)	50–3.0
Root mean-square deviation from ideal	
Bonds (Å)	0.0068
Angles (°)	1.092
Average B-factors (Å)^2	
Protein	69.2
DNA	127.0
Solvent	67.6

[a] $R_{merge} = \Sigma|I_h - \langle I_h \rangle|/\Sigma I_h$, where I_h is the mean of the measurements for a single hkl and R_{merge} is merging R-factor, which is the sum of the differences of all measurements from the average value of the measurement divided by the sum of all measurements.

[b] Residues included in each histone subunit: for H3, aa 38 to 135; for H3′, aa 38 to 135; for H4, aa 24 to 102; for H4′, aa 20 to 102; for macroH2A, aa 14 to 119; for macroH2A′, aa 14 to 119, for H2B, aa 30 to 22; for H2B′, aa 27 to 122. The remaining histone tails were too disordered to be included in the final model.

[c] R-factor = $\Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$, where R-factor is the overall agreement between amplitudes of two sets of structure factors, F_{obs} is the observed structure factor amplitude, and F_{calc} is calculated amplitude from the model.

To analyze histone deacetylase (HDAC)-chromatin interaction, nuclei of murine erythroleukemia cells were prepared and lysed in a buffer containing the following: 25 mM HEPES (pH 7.9), 10% glycerol, 600 mM KCl, 0.1% NP-40, 0.5 mM DTT, and 1 mM phenylmethylsulfonyl fluoride. The extracts were then sonicated (200 J) in order to generate chromatin fragments of about 400 bp, as shown by analysis of the DNA by electrophoresis on agarose gel. After centrifugation, immunoprecipitations using anti-HDAC1, -2, or -3 antibodies (Santa Cruz) were performed on the supernatants. Immunoprecipitated complexes were washed three times with the lysis buffer, eluted by denaturation in Laemmli buffer, and analyzed by Western blot analysis for the presence of macroH2A, HDAC1, and HDAC2.

Crystallographic coordinates and structure factors. The crystallographic coordinates and structure factors have been deposited in the Protein Data Bank under numbers 1U35 for the macro-NCP structure and 1YD9 for the nonhistone region of macroH2A.

RESULTS

macro-NCP differs from nonvariant NCP primarily in the L1 loop. The amino acid sequence of the histone domain of macroH2A (amino acids 1 to 122) is only 64% identical to major H2A from both *Xenopus laevis* and *Mus musculus* (Fig. 1A). There are two particular regions of sequence divergence that have the potential to cause structural and functional differences. First, the ladle-shaped H2A docking domain formed by H2A amino acids 83 through 108 is involved in an extensive interaction interface with one-half of the $(H3\text{-}H4)_2$ tetramer and guides the N-terminal helix of H3 (H3 αN) to interact with the penultimate 15 base pairs of the nucleosomal DNA. Second, the H2A L1 loops form the only interface between the two H2A-H2B dimers within a single NCP, perhaps facilitating the cooperative incorporation of the second H2A-H2B dimer in the NCP. This interface seemingly holds together the two gyres of the DNA superhelix and may thus play a role in regulating the dynamic behavior of nucleosomes.

We determined the crystal structure of the NCP reconstituted from mouse histones including the histone domain of macroH2A (amino acids 1 through 120; macro-NCP) to 3.0 Å resolution (Table 1). A representative region of the final electron density map is shown in Fig. 1B. The region of macroH2A-NCP shown differs between macroH2A and major H2A in residues 83, 84, and 88, which correspond to residues 80, 81, and 85 in macroH2A (L-I, Q-L, and R-A, respectively), and these differences in sequence were clearly visible in the initial electron density map. Despite the significant sequence differences between *Xenopus laevis* and mouse core histones in general and between *Xla*-H2A and macroH2A in particular, the structures of macro-NCP and *Xenopus laevis* NCP (*Xla*-NCP) are superimposable with a root mean-square deviation of <1Å. Figure 1C and 1D shows a superposition of the two structures in two different orientations. The path of the DNA is highly similar to that observed in major NCP (consistent with DNase footprinting experiments [1]); however, it shows significant differences in the way in which macro-NCP responds to the crystallization-induced stretching of the short half of their DNA (see references 21 and 29 for a general discussion of this phenomenon), which facilitates the compensation of the 1-base pair deficiency in length that is a direct consequence of the use of an even-numbered (146-bp) DNA fragment. While in all the nucleosome structures reported till now, the long and short halves of the DNA exhibit identical orientation throughout the crystal lattice, in the case of macro-NCPs, their relative orientation is indiscriminate. A convolution of these two orientations therefore results in electron density maps that show only 145 base pairs. The histone domain of macroH2A interacts with H2B in a manner very similar to *Xla*-H2A. Major interactions between the α2 helices of the two histones are maintained. There are no major differences in the protein-protein or protein-DNA interactions, with the notable exception of the L1-L1 interface between the two macroH2A moieties (Fig. 1C and D and 2A and B).

In *Xla*-NCP, the L1-L1 interface is stabilized by two intermolecular salt bridges between Glu41-Asn38′ and Glu41′-Asn38 (Fig. 2C). These interactions are replaced in macro-NCP by two sets of hydrophobic interactions. Lys37 is sandwiched between Tyr38 and Pro36′ and by interactions that are related via the axis of pseudosymmetry of the NCP, and Lys37′ is sandwiched between Tyr38′ and Pro36 (Fig. 2D). Together, these structural changes render the L1-L1 interface in macro-NCP less flexible and more hydrophobic, with potential effects on the stability of the histone octamer. In contrast, the interactions between the $macroH2A_{HD}$-H2B dimers and the $(H3\text{-}H4)_2$ tetramer remain unaffected by the numerous sequence variations in the macroH2A docking domain (Fig.

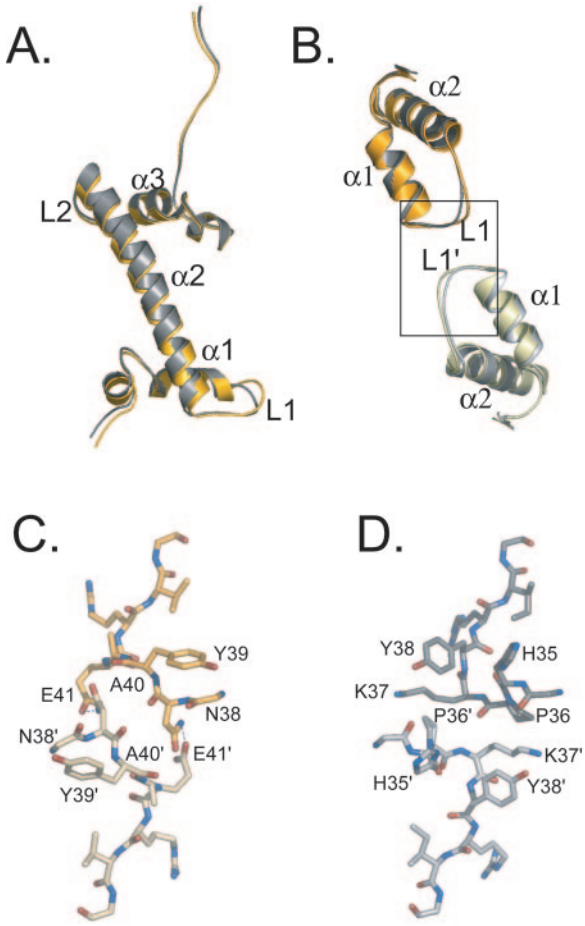

FIG. 2. The interface formed by two L1 loops differs significantly between macro-NCP and *Xla*-NCP. (A) Superposition of H2A and the histone domain of macroH2A in a view similar to that in Fig. 1C. macroH2A is shown in gray and major H2A in yellow. (B) Superposition of the L1 loops and the α-helices of macroH2A and macroH2A′ (gray and off-white, respectively) and of H2A and H2A′ (yellow and light yellow, respectively). Only minor changes in the path of the main chain of the L1 loop were observed. (C and D) Detailed view of the boxed area in panel B shows fundamental differences in the intermolecular interactions between two macroH2A chains (gray and off white) and two major H2A chains (yellow and light yellow) molecules, respectively.

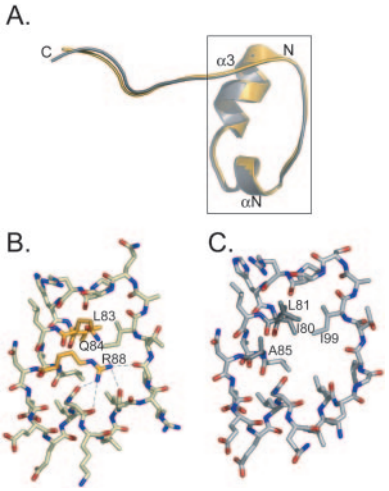

FIG. 3. The macroH2A docking domain is unaffected by sequence changes. (A) Superposition of the docking domains (amino acid residues 80 through 119) of macroH2A (light gray) and *Xla*-H2A (light yellow) demonstrates that there are no differences in the path of the main chain between the two docking domains. (B and C) Detailed view of the boxed area (amino acids 80 through 105 in majorH2A) in panel A highlights significant amino acid differences (L83 to I80, Q84 to L81, and in particular, R88 to A85) that may alter the stability of this area. The sequence differences are shown in dark gray (macroH2A) and dark yellow (*Xla*-H2A). Also shown are some of the hydrogen bonds formed by R88 to stabilize this domain in *Xla*-H2A.

1A). These substitutions do not result in structural changes in the histone main chain at the docking domain (Fig. 3A). In *Xla*-H2A, residues Leu83, Gln84, and, in particular, Arg88 appear to be essential in stabilizing the conformation of the docking domain in the absence of the $(H3\text{-}H4)_2$ tetramer (Fig. 3B). In macroH2A, these residues are changed to Ile80, Leu81, and Ala85 (Fig. 3C). The absence of the four hydrogen bonds that are being made by H2A R88 (Fig. 3B) is compensated for by a small hydrophobic interface between the unique macroH2A residues I80, L81, and I99 (Fig. 3C). Interestingly, the acidic side chains of H2A (E56, E61, E64, D90, E91, and E92) that form a pronounced acidic patch on the surface of the NCP are all conserved in macroH2A. This surface is essential for crystal contacts (16) and represents the most distinct surface characteristic that is present in all major-type NCPs.

Overall structure of the nonhistone domain of macroH2A (aa 180 through 367). To gain insight into the function of the nonhistone region of macroH2A, we determined the structure of the region from amino acids 180 to 367 to a resolution of 1.6 Å. Although the nonhistone region spans a region between amino acids 161 and 367, we used the region between amino acids 180 and 367 for crystallographic studies (Fig. 1A) because we observed degradation over time in the region between amino acids 161 and 180 during the crystallization of a fragment encompassing amino acids 161 to 367, indicating that this portion is at least partially disordered. Multiple-Wavelength Anomalous Dispersion phases were obtained from a gold-derivative to a resolution of 2.1 Å and were extended to 1.6 Å using a native data set. The structure was refined to a crystallographic R-factor of 23.5 (free R-factor = 26.0) (Table 2). A representative region of the final electron density map contoured at 2 σ is shown in Fig. 4A.

The domain falls into the α/β class of proteins with a seven-strand β-sheet and five α-helices (Fig. 4B and C). There are four molecules in the asymmetric unit. The region between amino acids 180 and 208, which has been predicted to be a leucine zipper based on sequence analysis (24), constitutes the first two strands of the seven-strand β-sheet (Fig. 4B and C). The β-sheet is protected on one face by α-helices, while the other face remains partially solvent exposed (Fig. 4B and C) and exhibits an extended hydrophobic region (Fig. 4D and E). This hydrophobic patch includes residues 182 to 185 (FTVL) and residues 355 to 359 (IGIYV). No regions of distinct surface charge were observed (Fig. 4D). The nonhistone region of macroH2A showed sequence homology with the widely found A1pp domain, which was found in a wide variety of other

TABLE 2. Data collection and refinement statistics for the MAD experiment

Statistic	Value(s)[a]
Space group	$P2_12_12_1$; $P2_12_12_1$; $P2_12_12_1$
Wavelength (λ) (Å)	1.074800; 1.039990; 1.074853
Resolution range (Å)	50–2.1; 50–2.1; 50–1.6
No. of unique reflections	42,574; 42,686; 93,213
% R_{merge}[b] (overall/last shell)	6.7/34.9; 7.4/37.7; 6.0/41.0
Resolution range	50–1.6
Cell dimensions	a = 83.11, b = 89.79, c = 95.68, α = 90, β = 90, γ = 90
No. of reflections (working/test)	88,260/2,757
Map correlation coefficient[c]	0.7
No. of protein residues	750
No. of water molecules	540
No. of molecules in asymmetric unit	4
R-factor[d]/R_{free}	0.235/0.260
B-factor average	
Protein	30.94
Solvent	39.36
RMSD from ideality	
Bonds (Å)	0.0048
Angles (°)	1.2613
Ramachandran plot (% in allowed region)	93

[a] Values for the space group, wavelength, resolution range, number of unique reflections, and R_{merge} are given in three distinct groups: peak, high-energy remote, and native, respectively.

[b] $R_{merge} = \Sigma|I_h - \langle I_h \rangle|/\Sigma I_h$, where I_h is the mean of the measurements for a single hkl.

[c] Correlation coefficient = $\int_v \rho_1(x)\rho_2(x)dx/[\int_v \rho_1(x)^2dx \int_v \rho_2(x)^2dx]^{1/2}$ (26).

[d] R-factor = $\Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$.

proteins and biological contexts. Additionally, it exhibited structural similarity with proteins without discernible sequence homology. For example, the regulatory domain of leucine aminopeptidase has the highest degree of structural similarity (Dali Z-score, 10.8), yet its sequence is only ~11% homologous. The wide variety of species and biological contexts in which this domain is found strongly suggests that this fold is functionally very versatile and allows much room for speculation about its function.

The nonhistone domain of macroH2A interacts with HDAC1. macroH2A1.2 is found in high concentration on the inactive X-chromosome of female mammals, which resembles heterochromatin, and is also found to be generally repressive to transcription (7, 25). We hypothesized that macroH2A may recruit one or more activities that may be responsible for the altered state of covalent modifications on heterochromatin, such as HDACs, as well as repressive histone lysine methyltransferases. We independently expressed three class I HDACs, HDAC1, HDAC2, and HDAC3, and one histone lysine methyltransferase, Suv39h1, together with the nonhistone region of macroH2A (aa 160 through 370). We found that both HDAC1 and HDAC2 but neither HDAC3 nor Suv39h1 efficiently interacted with the nonhistone region of macroH2A (Fig. 5A). In this experiment, Cdyl (chromodomain-Y-like), previously shown to interact with HDAC1/2 and Suv39h1 (6; unpublished results), was used as a positive control. HDAC1 and HDAC2 are found together in major HDAC complexes and are functionally distinct from HDAC3. These data confirm these functional distinctions.

We created deletion and point mutants to determine the regions that are important for the association between the nonhistone region of macroH2A and HDAC1 (Fig. 5B). Both the N-terminal (residues 182 to 185 –FTVL) and C-terminal (residues 358 to 362; IGIYV corresponding to and identical to residues 355 to 359 in macroH2A1.1) amino acid sequences that contribute to an extended hydrophobic patch on the surface of the nonhistone region participate in stabilizing macroH2A-HDAC1 interaction (Fig. 5C and D). IP experiments show that the replacement of both regions with charged residues affects the stability of the interaction and partially relieves the repressor activity of the domain. However, in vivo tests (transcriptional repression) showed that contribution of the C-terminal hydrophobic residues is critical in the repressor activity of the nonhistone region, since the m2 mutant showed the same repressor activity as the double mutant m1/2 (Fig. 6A). These data suggest that although the m2 mutant interacts efficiently with HDAC1, there are enough structural alterations to affect the repressor activity of the domain. We also found that the interaction depends in part on the region between amino acids 160 and 180 which is not included in the present structure. The deletion of this region significantly affects the stability of macrodomain-HDAC1 interactions (Fig. 5D) and completely abolishes the repressive activity of this domain (Fig. 6A, S-Wt). Importantly, we see a potential connection between the macroH2A-mediated recruitment of HDACs and the hypoacetylation of chromatin. Indeed, the immunoprecipitation of Ha-tagged macroH2A-containing chromatin showed that the presence of macroH2A in oligonucleosomes coincides with a hypoacetylated state of H3 in the region (Fig. 6B and C). We confirmed that similar quantities of nucleosomes and histones have been immunoprecipitated after analyzing a fraction of the immunoprecipitated materials on a sodium dodecyl sulfate-polyacrylamide electrophoresis gel (Fig. 6C, silver-stained panel and input panel of the Western blot analysis). These same materials were then used to test H3 acetylation. This blot was then controlled by anti-H3 and anti-H2A antibody, which showed that the loading and transfer were correct.

In order to confirm the association between macroH2A-containing nucleosomes and HDAC1/2, oligonucleosomes generated after the sonication of murine erythroleukemia cells were immunoprecipitated using anti-HDAC1, -2, or -3 antibodies on the supernatants and the presence of macroH2A and HDACs was monitored by Western blotting (Fig. 6E). Please note that HDAC3 migrates at the same position as the immunoglobulin (Ig) heavy chain band and that the immunoprecipitated protein could therefore not be visualized by Western blotting (data not shown). We found that macroH2A coimmunoprecipitates with endogenous HDAC1 and HDAC2 (Fig. 6E, upper panel). This strongly indicates that macroH2A and HDAC1/2 coexist in chromatin under physiological concentra-

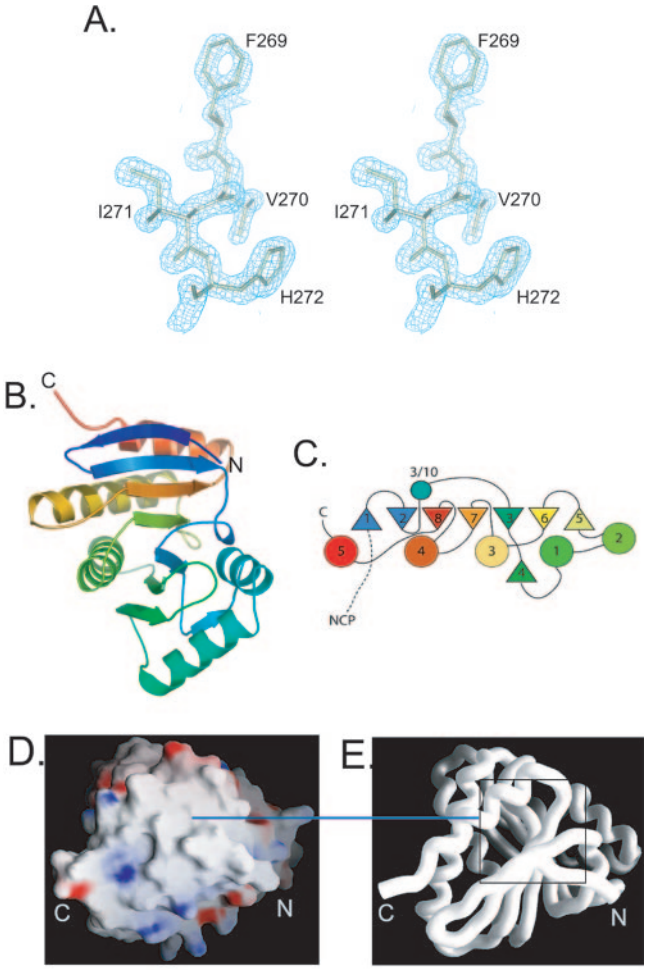

FIG. 4. Structure of the nonhistone region of macroH2A. (A) Stereo view of a section of the 2Fo-Fc electron density map, calculated at 1.6 Å and contoured at 2 σ, clearly showing a part of the sequence of the nonhistone region (F269-H272). (B) Overall structure of the nonhistone region (aa 180 through 370). The N terminus (N) is in blue and the C terminus (C) in red with a gradient of the colors of the visible spectrum in between. (C) A schematic representation of the fold in Fig. 1B. Beta-strands are depicted as arrowheads and helices as circles. (D) The surface representation of the nonhistone region. Basic regions are in blue, acidic regions in red, and neutral regions in white. (E) The C-α trace of the nonhistone region of macroH2A in exactly the same orientation as in panel D. The boxed area encompasses a large hydrophobic region that includes residues F183 to L186 and I356 to V360.

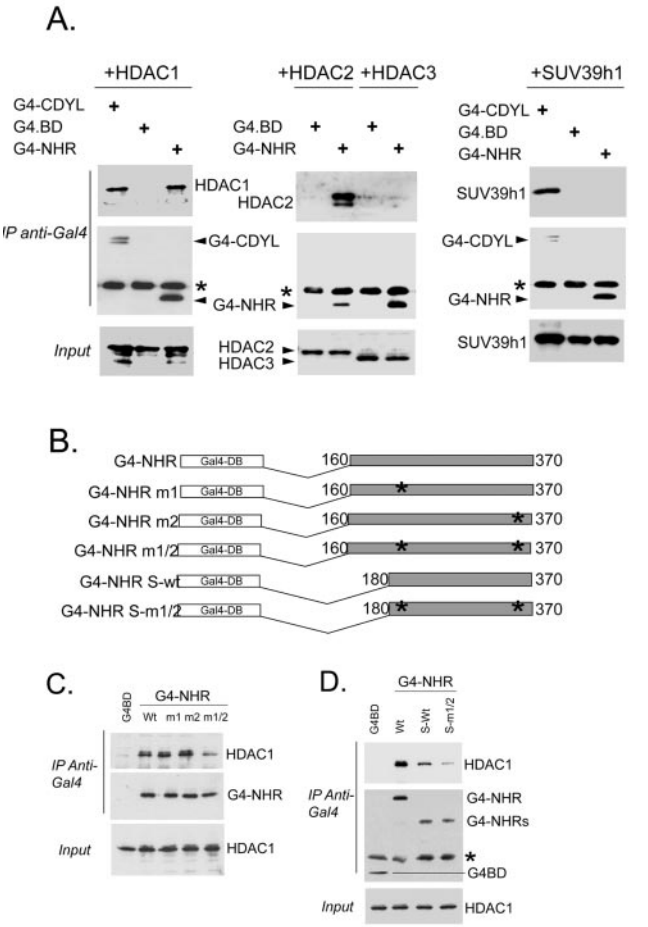

FIG. 5. The nonhistone region (NHR) of macroH2A1.2 is associated with HDAC1. (A) Cos cells were cotransfected with 3 μg of expression vectors for Flag-tagged HDAC1, -2, or -3 or myc-tagged SUV39h1, and 3 μg of vectors expressing the NHR of macroH2A (aa 160 through 370) were fused to the DNA-binding domain of Gal4 (G4-NHR). The Gal4 binding domain alone (G4.BD) and G4-Cdyl were used as negative and positive controls, respectively. The Gal4 fusion proteins were immunoprecipitated from whole-cell extracts with 1 μg of an anti-Gal4 antibody, and proteins present in the complexes were analyzed with anti-Flag (recognizing HDAC1, HDAC2, and HDAC3) and anti-myc (recognizing SUV39h1) antibodies (upper panel). The same blot was then probed with an anti-Gal4 antibody (middle panel). G4.DB is not visible in the range of molecular weight shown (Fig. 4D). Asterisks indicate the Ig-heavy chains. The lower panels show the amounts of HDACs or SUV39h1 present in 5% of each input. (B) Parts of the NHR involved in the interaction with HDAC1. Schematic representations of G4-NHR constructs m1 and m2 have been generated by replacing 2 amino acids of hydrophobic patches (F182 and T183 in m1 and I360 and Y361 in m2) with arginines. The double mutant (m1/2) corresponds to the association of m1 and m2 mutations. (C and D) Coimmunoprecipitations of Flag-HDAC1 with G4-NHR proteins were performed as described in the legend to Fig. 1A, except that lysis of cells and washing of immunoprecipitated complexes were performed under more stringent conditions. Asterisks indicate the Ig-light chains. Wt, wild type.

tions of the two proteins. Our study therefore suggests the possibility that the nonhistone region of macroH2A may play a vital role in heterochromatin formation by recruiting class I HDACs.

DISCUSSION

We have shown that structural and biochemical differences between NCPs containing the histone-like domain of macroH2A and those reconstituted with major-type H2A can largely be reduced to a 4-amino-acid region in the L1 loop that connects two α-helices of the histone fold domain of H2A or macroH2A. The nature of this interface is completely altered in macro-NCP, and these changes are responsible for a less stringent requirement for high ionic strength to hold the histone octamer together in the absence of DNA (S. Chakravarthy and K. Luger, manuscript in preparation). An altered

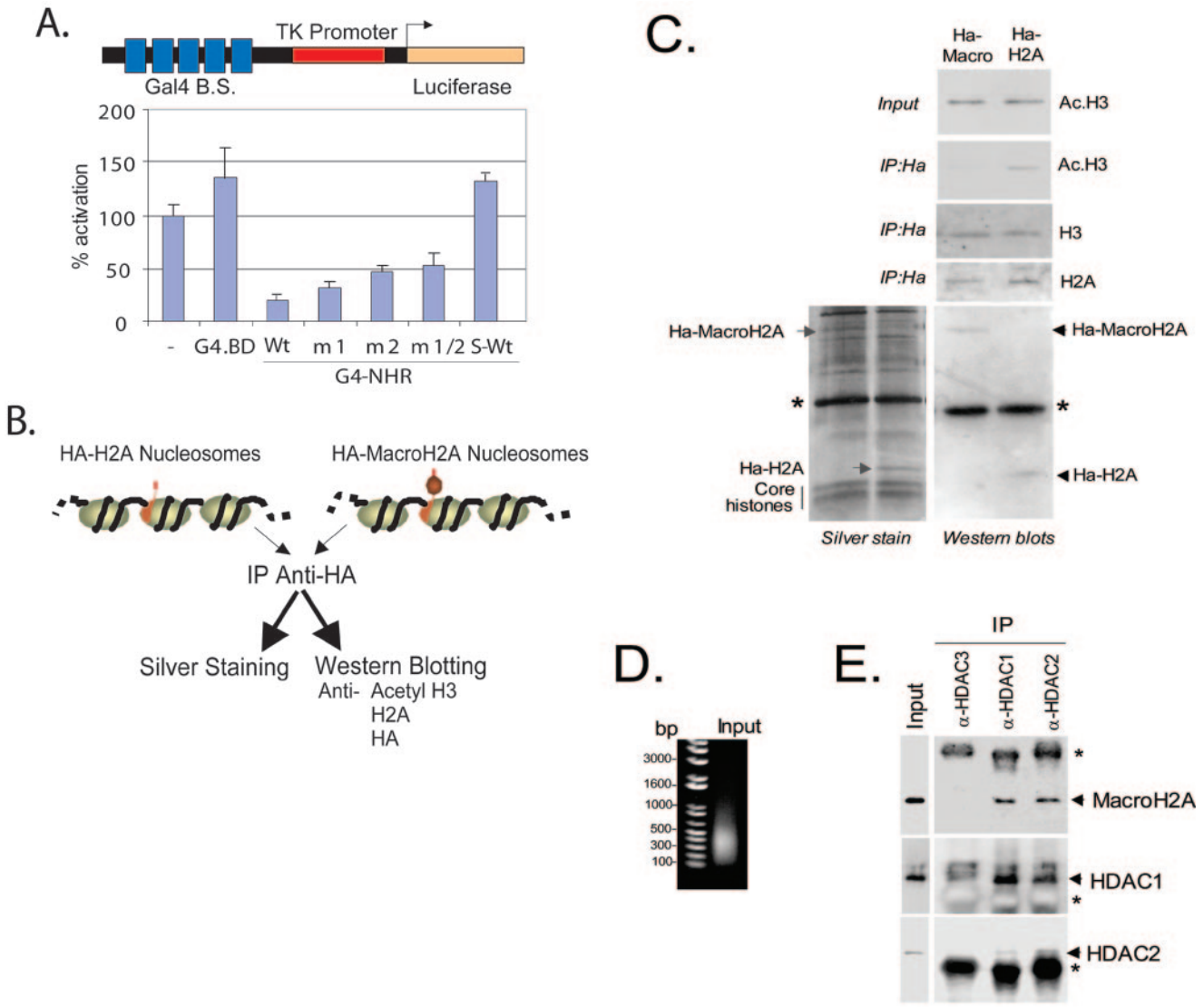

FIG. 6. The nonhistone region (NHR) of macroH2A1.2 is associated with hypoacetylated chromatin. (A) A total of 0.5 μg of each G4-NHR construct [or Gal4.DB or an empty vector (−) as controls] was cotransfected in Cos cells with 1 μg of a luciferase plasmid reporter containing five Gal4 sites in its promoter. In each transfection, 100 ng of pCMV-β-gal control plasmid was also used for normalization purposes. Luciferase activity was measured on cell extracts 24 h after transfection and normalized to that of β-galactosidase. Mean values of at least four independent assays are represented. Wt, wild type. (B) Schematic representation of the chromatin immunoprecipitation assay. Cos cells tranfected with expression vectors of either Ha-tagged H2A or Ha-tagged macroH2A in the presence of the ectopically expressed HAT p300 (in order to enhance background H3 acetylation to visualize HDAC activity) were lysed, and chromatin was fragmented by sonication in order to obtain DNA fragments with a mean size of 600 to 1,000 bp. Chromatin fragments containing nucleosomes with Ha-H2A or Ha-macroH2A were immunoprecipitated with an anti-Ha antibody and analyzed by silver staining or Western blotting for the presence of histones and the level of their acetylation. (C) Immunoprecipitated Ha-macroH2A and Ha-H2A were detected by anti-Ha Western blotting (lower right panel) and were also seen on a silver-stained gel (left panel, indicated with arrows). Acetylation level of coprecipitated histones was analyzed with antibodies recognizing acetylated histone H3 (third panel on the right) and compared to 5% of the input (upper right panel). Asterisks indicate the Ig light chains. The same blot was also probed with anti-H2A and anti-H3 antibodies (as indicated). (D and E) Nuclei of murine erythroleukemia cells were lysed and sonicated in order to generate small soluble chromatin fragments. (D) DNA analysis on an agarose electrophoresis gel showing that the fragments have a mean size of about 400 bp. (E) HDAC1, -2, or -3 was immunoprecipitated from the extracts using the specific respective antibodies (HDAC1, -2, or -3 IP), and the presence of macroH2A in the complex was detected by an anti-macroH2A Western blot (upper panel). The HDAC1 and -2 panels show the corresponding immunoprecipitated proteins as controls. Input corresponds to 5% of the extract. Asterisks indicate the Ig-heavy chains.

L1-L1 interface may also stabilize the two gyres of DNA in the nucleosome and may in turn cause an increased resistance to the nucleosome remodeling machinery (SWI/SNF for instance, as suggested in reference 4). Recent studies have shown that transcription by RNA polymerase II involves the removal of one H2A-H2B dimer from the nucleosome (5, 13). While the H2A-H2B dimer is stabilized by the L1-L1 interface and interactions with the $(H3\text{-}H4)_2$ tetramer, it is reasonable to expect alterations in the L1-L1 interface to preclude the removal of an H2A-H2B dimer. In other words, a conformationally

inflexible L1-L1 interface may make the macro-NCP more refractory to the machinations of the transcriptional machinery. In vivo targeting studies, which were guided by results obtained from biochemical and structural analysis, showed that 4 amino acids in the L1 loop are to a large extent responsible for the in vivo targeting of macroH2A to the inactive X-chromosome (Dmitri Nusinov and Barbara Panning, personal communication). This is reminiscent of the situation with H3.3, where single amino acid changes between major H3 and H3.3 are sufficient to channel these histones to replication-dependent and replication-independent assembly pathways, respectively (2). It is likely that specific chromatin assembly and exchange factors, like those found for histone H2A.Z (14, 15, 20) and H3.3 (30), will be found for other histone variants, including macroH2A, and it will be of interest to see whether the regions pinpointed in this study are responsible for targeting specificity.

The structure of the nonhistone region of macroH2A (aa 180 through 370) exhibits a compact α/β fold without extending surface loops and a rather nondescript surface charge distribution including a large hydrophobic patch. Based on in vitro results, we propose that this domain may be involved in the recruitment of HDACs. This conclusion is further corroborated by the fact that H3 was found hypoacetylated in chromatin regions containing macroH2A nucleosomes. These observations are compatible with recent work that shows that macroH2A is localized to hypoacetylated regions of chromatin, e.g., the inactive X-chromosome (7) and the pericentric heterochromatin (10). The recruitment of HDACs is correlated with the hypoacetylation of chromatin and repression of transcription of a reporter gene. Mutating the hydrophobic residues to charged residues partially relieves this repression. It has also been suggested that the macrodomain may have an enzymatic function based on sequence homology with a diverse group of proteins containing the A1pp domain (from the Appr-1″-p processing enzyme YBR022Wp in *Saccharomyces cerevisiae*) (19). While this has not been confirmed experimentally for macroH2A, the notion is supported indirectly by recent results showing that the macrodomain in a variety of proteins (including macroH2A) has a high affinity for monomeric and polymeric ADP-ribose, a known intermediate in several vital cellular pathways (12). Structural studies with the archaeal protein AF1521 showed a fold remarkably similar to that of macrodomain shown in this study and also a similarity with the P-loop containing nucleoside triphosphate hydrolases (3). Also remarkable is the degree of sequence conservation seen in regions that were determined to be important in ADP-ribose binding (12). It will be interesting to see if ADP-ribose binding has any influence on the ability of macroH2A to act as a recruitment platform for other chromatin-associated proteins such as HDAC. Site-directed mutagenesis of the conserved regions that facilitate ADP-ribose binding should answer some of these questions in the near future. In spite of the wide range of sequence and structural similarities, it is important to keep in mind the possibility of an entirely novel *modus operandi* for macroH2A.

Together, our data suggest that the incorporation of macroH2A has two effects on chromatin. First, it may locally affect nucleosome structure and stability as well as the propensity of macroH2A containing chromatin to be remodeled through specific sequence changes in the histone domain, most notably in the L1 loop. Second, the nonhistone region may be involved in the recruitment of nonhistone-chromatin-associated proteins such as HDACs, thereby altering the acetylation state of histone tails, with possible effects on higher order structure.

ACKNOWLEDGMENTS

We thank Danny Rangasamy, David Tremethick, Thomas Jenuwein, and Ed Seto, who kindly donated expression clones for mouse H2B and H3, Flag-HDAC1, -2, and-3, and myc-Suv39h1. We also thank Pam Dyer for invaluable help with reagents and Sandrine Curtet-Benitski and Edwige Col for technical assistance. We are grateful to B. Panning and D. Nusinov for allowing us to cite their unpublished data.

This work was supported by research grants from the March of Dimes Birth Defects Organization and the NIH to K.L. and from Cancéropôle Rhône-Aples to S.K.

REFERENCES

1. **Abbott, D. W., V. S. Ivanova, X. Wang, W. M. Bonner, and J. Ausio.** 2001. Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation. J. Biol. Chem. **276:**41945–41949.
2. **Ahmad, K., and S. Henikoff.** 2002. The histone variant H3.3 marks active chromatin by replication- independent nucleosome assembly. Mol. Cell **9:**1191–1200.
3. **Allen, M. D., A. M. Buckle, S. C. Cordell, J. Lowe, and M. Bycroft.** 2003. The crystal structure of AF1521 a protein from Archaeoglobus fulgidus with homology to the non-histone domain of macroH2A. J. Mol. Biol. **330:**503–511.
4. **Angelov, D., A. Molla, P. Y. Perche, F. Hans, J. Cote, S. Khochbin, P. Bouvet, and S. Dimitrov.** 2003. The histone variant MacroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. Mol. Cell **11:**1033–1041.
5. **Belotserkovskaya, R., S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg.** 2003. FACT facilitates transcription-dependent nucleosome alteration. Science **301:**1090–1093.
6. **Caron, C., C. Pivot-Pajot, L. A. van Grunsven, E. Col, C. Lestrat, S. Rousseaux, and S. Khochbin.** 2003. Cdyl: a new transcriptional co-repressor. EMBO Rep. **4:**877–882.
7. **Costanzi, C., and J. R. Pehrson.** 1998. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature **393:**599–601.
8. **Dyer, P. N., R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan, and K. Luger.** 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. **375:**23–44.
9. **Fan, J. Y., F. Gordon, K. Luger, J. C. Hansen, and D. J. Tremethick.** 2002. The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. Nat. Struct. Biol. **19:**172–176.
10. **Grigoryev, S. A., T. Nikitina, J. R. Pehrson, P. B. Singh, and C. L. Woodcock.** 2004. Dynamic relocation of epigenetic chromatin markers reveals an active role of constitutive heterochromatin in the transition from proliferation to quiescence. J. Cell Sci. **117:**6153–6162.
11. **Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjelgaard.** 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. Sect. A **47:**110–119.
12. **Karras, G. I., G. Kustatscher, H. R. Buhecha, M. D. Allen, C. Pugieux, F. Sait, M. Bycroft, and A. G. Ladurner.** 2005. The macro domain is an ADP-ribose binding module. EMBO J. **24:**1911–1920.
13. **Kireeva, M. L., W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, and V. M. Studitsky.** 2002. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. Mol. Cell **9:**541–552.
14. **Kobor, M. S., S. Venkatasubrahmanyam, M. D. Meneghini, J. W. Gin, J. L. Jennings, A. J. Link, H. D. Madhani, and J. Rine.** 2004. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A. Z into euchromatin. PLoS Biol. **2:**E131.
15. **Krogan, N. J., M. C. Keogh, N. Datta, C. Sawa, O. W. Ryan, H. Ding, R. A. Haw, J. Pootoolal, A. Tong, V. Canadien, D. P. Richards, X. Wu, A. Emili, T. R. Hughes, S. Buratowski, and J. F. Greenblatt.** 2003. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol. Cell **12:**1565–1576.
16. **Luger, K., A. W. Maeder, R. K. Richmond, D. F. Sargent, and T. J. Richmond.** 1997. X-ray structure of the nucleosome core particle at 2.8 Å resolution. Nature **389:**251–259.
17. **Luger, K., T. J. Rechsteiner, and T. J. Richmond.** 1999. Preparation of nucleosome core particle from recombinant histones. Methods Enzymol. **304:**3–19.

18. **Malik, H. S., and S. Henikoff.** 2003. Phylogenomics of the nucleosome. Nat. Struct. Biol. **10:**882–891.
19. **Martzen, M. R., S. M. McCraith, S. L. Spinelli, F. M. Torres, S. Fields, E. J. Grayhack, and E. M. Phizicky.** 1999. A biochemical genomics approach for identifying genes by the activity of their products. Science **286:**1153–1155.
20. **Mizuguchi, G., X. Shen, J. Landry, W. H. Wu, S. Sen, and C. Wu.** 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science **303:**343–348.
21. **Muthurajan, U. M., Y. J. Park, R. S. Edayathumangalam, R. K. Suto, S. Chakravarthy, P. N. Dyer, and K. Luger.** 2003. Structure and dynamics of nucleosomal DNA. Biopolymers **68:**547–556.
22. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode, p. 307–326. *In* J. N. Abelson, M. I. Simon, Jr., C. W. Carter, and R. M. Sweet (ed.), Methods in enzymology, vol. 276: macromolecular crystallography, part A. Academic Press, New York, N.Y.
23. **Pehrson, J. R., and V. A. Fried.** 1992. MacroH2A, a core histone containing a large nonhistone region. Science **257:**1398–1400.
24. **Pehrson, J. R., and R. N. Fuji.** 1998. Evolutionary conservation of histone macroH2A subtypes and domains. Nucleic Acids Res. **26:**2837–2842.
25. **Perche, P., C. Vourc'h, L. Konecny, C. Souchier, M. Robert-Nicoud, S. Dimitrov, and S. Khochbin.** 2000. Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. Curr. Biol. **10:**1531–1534.
26. **Read, R. Y.** 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr. Sect. A **42:**140–149.
27. **Rice, L. M., Y. Shamoo, and A. T. Brunger.** 1998. Phase improvement by moluti-start simulated annealing refinement and structure-factor averaging. J. Appl. Crystallogr. **31:**798–805.
28. **Suto, R. K., M. J. Clarkson, D. J. Tremethick, and K. Luger.** 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat. Struct. Biol. **7:**1121–1124.
29. **Suto, R. K., R. S. Edayathumangalam, C. L. White, C. Melander, J. M. Gottesfeld, P. B. Dervan, and K. Luger.** 2003. Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. J. Mol. Biol. **326:**371–380.
30. **Tagami, H., D. Ray-Gallet, G. Almouzni, and Y. Nakatani.** 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell **116:**51–61.
31. **Terwilliger, T. C.** 2003. SOLVE and RESOLVE: automated structure solution and density modification. Methods Enzymol. **374:**22–37.