The preferential binding of histone H1 to DNA scaffold-associated regions is determined by its C-terminal domain

Alicia Roque, Mary Orrego, Imma Ponte and Pedro Suau*

Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

ABSTRACT

Histone H1 preferentially binds and aggregates scaffold-associated regions (SARs) via the numerous homopolymeric oligo(dA).oligo(dT) tracts present within these sequences. Here we show that the mammalian somatic subtypes H1a,b,c,d,e and H1° and the male germline-specific subtype H1t, all preferentially bind to the *Drosophila* histone SAR. Experiments with the isolated domains show that whilst the C-terminal domain maintains strong and preferential binding, the N-terminal and globular domains show weak binding and poor specificity for the SAR. The preferential binding of SAR by the H1 molecule thus appears to be determined by its highly basic C-terminal domain. Salmine, a typical fish protamine, which could have its evolutionary origin in histone H1, also shows preferential binding to the SAR. The interaction of distamycin, a minor groove binder with high affinity for homopolymeric oligo(dA).oligo(dT) tracts, abolishes preferential binding of the C-terminal domain of histone H1 and protamine to the SAR, suggesting the involvement of the DNA minor groove in the interaction.

INTRODUCTION

H1 linker histones are thought to be primarily responsible for the condensation of the nucleosome chain in the thick chromatin fibre. It is currently accepted that histone H1 could have a regulatory role in transcription through the modulation of chromatin higher-order structure. H1 has been described as a general transcriptional repressor because it contributes to chromatin condensation, which limits the access of the transcriptional machinery to DNA. Other studies indicate that H1 may regulate transcription at a more specific level, participating in complexes that either activate or repress specific genes (1–8). Preferential binding to SARs (for scaffold-associated regions; also termed matrix-associated regions, MARs) (9) and participation in nucleosome positioning (10) are other mechanisms by which H1 could contribute to transcriptional regulation.

H1 has multiple isoforms. In mammals, six somatic subtypes, designated H1a–e and H1°, a male germline-specific subtype, H1t, and an oocyte-specific subtype, H1oo, have been identified (11–14). The subtypes differ in their timing of expression (15), extent of phosphorylation (16) and turnover rate (17,18). Analysis of the evolution of vertebrate H1 subtypes has shown that amino acid substitution rates differ among subtypes by almost one order of magnitude, suggesting that each subtype has acquired a unique function (19). Differences in DNA condensing capacity have also been demonstrated for some subtypes (20–22).

Histone H1 has been characterized as an SAR-binding protein (9). SARs were identified as DNA sequences rich in AT base pairs (>70%) and in homopolymeric oligo(dA).oligo(dT) tracts (A-tracts) (23) that were specifically bound by the nuclear and the metaphase scaffold. SARs have been proposed as DNA elements that would define the bases of chromatin loops in eukaryotic cells (23), and that could also be involved in chromosome dynamics (24). In addition to H1, several SAR-binding proteins have been identified, including topoisomerase II, lamin B1, nucleoline, HMG I/Y, SAT B1, SAF-A and SAF-B (25–30). It has been proposed that regulated H1 dissociation or assembly with SARs is implicated in the regional opening or closing of chromatin loops and, consequently, contributes to transcriptional regulation (9). The high affinity cooperative binding of H1 and other SAR-binding proteins to SARs is determined by the presence of the A-tracts, rather than a precise base sequence (25,31). A-tracts have peculiar structural properties, including a narrower minor groove (32,33). The importance of the conformational features of the DNA in H1 binding is stressed by the preferential affinity of H1 for supercoiled DNA, cruciforms and DNA fragments with intrinsic curvature (34–36).

H1 linker histones present a tripartite structure consisting of a central globular domain flanked by highly basic N- and C-terminal tail-like domains (37). The N- and C-terminal domains are very different in length, the C-terminal tail comprising almost 50% of the protein. The distribution of charge in the C-terminal domain is extremely uniform in spite of the variation in sequence of the different H1 subtypes (38). The structure of the globular domain has been described by X-ray diffraction (39) and nuclear magnetic resonance (40,41). It contains a three-helix bundle, which resembles the winged-helix motif found in some sequence-specific DNA-binding

*To whom correspondence should be addressed. Tel: +34 93 5811391; Fax: +34 93 5811264; Email: pere.suau@uab.es

proteins, and a C-terminal β-sheet. The N- and C-terminal domains are largely unstructured in solution. However, they acquire a substantial amount of secondary structure upon interaction with DNA. Helix and turn elements, inducible upon interaction with the DNA, have been described in both the N- and C-terminal domains (42–45). The distinct structure of the H1 domains suggests that they could play specific functions in chromatin structure. The N-terminal domain could contribute to the binding stability of the globular domain; the globular domain is very likely to localize the molecule in the nucleosome; while the C-terminal domain is the main region of the molecule involved in chromatin condensation through binding and neutralization of the charge of the linker DNA.

The tripartite structure of histone H1 raises the question of the involvement of the individual H1 domains in the preferential binding of the protein to SARs. Here, we identify the C-terminal domain as the main determinant of the SAR-binding properties of histone H1. We also show that salmine, a typical protamine that could have its evolutionary origin in histone H1 (46), preferentially binds to SARs. The possible significance of these findings in relation to H1 and SAR functions is discussed.

MATERIALS AND METHODS

Peptides and protamine

The peptides Ac-EKTPVKKKARKAAGGAKRKTSG-NH$_2$ (NE-1) and Ac-TENSTSAPAAKPKRAKASKK-NH$_2$ (NH-1) were synthesized by standard methods (NE-1 by Neosystem Laboratoire, Strasbourg, France, and NH-1 by DiverDrugs, Barcelona, Spain). Peptide homogeneity was determined by HPLC on Kromasil C8. The peptide composition was confirmed by amino acid analysis and the molecular mass was checked by mass spectrometry. The sequence of NE-1 corresponds to residues 15–36 at the N-terminus of mouse H1e and that of NH-1 to residues 1–20 of mouse H1°. Protamine (salmine) was from salmon (Sigma P4005).

Separation of histone H1 subtypes

Nuclei were isolated from the brain of adult mice (47,48). Histone H1 was extracted with 0.35 M NaCl, following the method of García-Ramírez *et al.* (49). The mixture of subtypes was digested with alkaline phosphatase to eliminate small amounts of phosphorylated forms that could be present. H1° was purified by gel-filtration chromatography, according to Böhm *et al.* (50). The subtypes H1a–e were separated by reverse phase HPLC according to Brown *et al.* (51). All subtypes were obtained as homogeneous peaks, except H1d and H1e, which largely overlapped. The latter subtypes were separated by acetic acid/urea gel electrophoresis (52) and recovered by electroelution using a Biotrap camera (Schleider & Schuell). H1t was purified from mouse testes according to Khadake *et al.* (21). Before being used in binding experiments, H1 subtypes were subjected to a cycle of denaturation/renaturation by stepwise dialysis from 6 M urea into, successively, 3.0 M, 1.5 M, 0.7 M, 0.3 M and 0.0 M urea, in 0.2 M NaCl, 0.01 M phosphate buffer, pH 7. Finally, the proteins were dialysed against 0.14 M NaCl, 0.01 M phosphate buffer,

pH 7.0. The concentration of protein was estimated by amino acid analysis.

Cloning, expression and purification of the globular and C-terminal domains of histone H1 subtypes

The sequences encoding the globular domains of histones H1° and H5, and the C-terminal domains of histones H1e, H1° and H1t were cloned and expressed. All gene fragments were amplified from mouse genomic DNA by PCR. The primers were 5′-GGCCGCCCATATGTCCACGGACCACCCCAAG-3′ and 5′-CTTGGATCCCTACGACCTCTTGGGCTC-3′ for the globular domain of H1°; 5′-GGCATCGCATAT-GTCGGCATCGCACCCCACC-3′ and 5′-GCCGGATCCT-TAGGACCTCTTGGCC-3′ for the globular domain of H5; 5′-CCACCATGGATGAGCCTAAAAGGTC-3′ and 5′-GGA-GATCTCTTCTTCTTGCTGGCCCTCT-3′ for the C-terminal domain of H1°; 5′-AAACCATGGCTGCTTCCGGTGAGGC-TAA-3′ and 5′-ACAGATCTCTTTTTCTTGGCTGCGGT-TTT-3′ for the C-terminal domain of H1e; and 5′-GTA-CCATGGCGGCTTCAGGGAACGAC-3′ and 5′-ACGGA-TCCCTTCCTCCCTGCTGCCTTCCT-3′ for the C-terminal domain of H1t. The amplification products of H1° and H5 globular domains were cloned in the pET11b vector (Novagen), using NdeI and BamHI restriction sites to yield the expression vectors pGH1° and pGH5, respectively. The C-terminal domains were cloned in the pQE-60 vector (Qiagen) using the NcoI and BglII restriction sites to yield the expression vectors pCTH1°, pCTH1e and pCTH1t.

The recombinant plasmids pGH1° and pGH5 were transformed into *E.coli* BL21(DE3). Cells were grown to an OD$_{600}$ of 0.8 and then induced with 1 mM IPTG, allowing expression to proceed for 4 h at 37°C. Cells were then harvested and stored at −80°C. The protein was purified according to the protocol described for the globular domain of chicken histone H5 (53).

The expression vectors pCTH1°, pCTH1e and pCTH1t were transformed into *E.coli* M15 (Qiagen). Cells were grown and induced as previously described for pGH1° and pGH5. Cells were lysed in lysis buffer (0.05 M NaH$_2$PO$_4$, 0.75 M NaCl, 0.02 M imidazol) plus 4 M guanidine hydrochloride, pH 8.0 for 15 min at room temperature. Guanidine hydrochloride was found to be necessary to avoid degradation and aggregation of the expressed protein. The extract was centrifuged at 20 000 *g* for 25 min. The supernatants were loaded on a HiTrap chelating HP column (Amersham Biosciences) equilibrated with lysis buffer. The column was then washed in three steps with lysis buffer containing increasing amounts of imidazol: 40, 60 and 80 mM. Finally, the proteins were eluted with 250 mM imidazol in lysis buffer and desalted by gel filtration through Sephadex G-25 (Amersham Biosciences).

Preparation of DNA fragments

An SAR fragment of 657 bp from the histone cluster of *Drosophila melanogaster* was obtained by digestion of p1314 (9) with KpnI and BamHI. Another DNA fragment, of 587 bp, was excised from pUC19 by digestion with HaeIII. Both inserts were separated on agarose gels and electroeluted using a Biotrap camera (Schleider & Schuell). The 587 bp pUC19 fragment was extended from both ends to obtain a fragment of 763 bp. The long fragment was obtained by PCR

on pUC19 using the primers 5′-GCGGTTAGCTCCTTCG-GTCCTC-3′ and 5′-CACCCGCTGACGCGCCCTGACG-3′. An AT-rich sequence (75% AT) was prepared by polymerization of the 5′-phosphorylated oligomer.

5′-CTATGATATATAGATAGTTAATGTAATATGATA-TAGATATAGGGATCC-3′, annealed with a complementary sequence that left five overhanging nucleotides. The annealed DNA was ligated overnight at 16°C with T4-ligase (Roche). The products of the ligation reaction were separated on agarose gels. Fragments ranging in size from ∼500 to 5000 bp were electroeluted using a Biotrap camera (Scheider & Schuell).

Binding assays

Binding experiments were performed by mixing equivalent amounts of SAR and pUC19 fragments with the proteins at different ratios. Binding conditions were 0.01 M phosphate buffer, pH 7.0, 5% glycerol and either 0.07 or 0.14 M NaCl. After 1 h of incubation at 37°C, the mixtures were centrifuged at 14 000 g for 10 min. The pellets and the supernatants were digested with proteinase K at 37°C overnight and the DNA was purified by phenol–chloroform extraction. The proportions of the SAR and pUC19 fragments in the complex and the supernatant were analysed by electrophoresis on 2% agarose (Metaphor) gels.

In some experiments, histone H1 C-terminal domain and protamine binding was performed in the presence of the DNA-binding drug distamycin A (Sigma). The DNA was preincubated with the drug for 30 min at 37°C before adding the proteins. The mixture was then incubated for 90 min at 37°C, as described.

RESULTS

All histone subtypes are SAR-binding proteins

The specific interaction of H1 with SARs was originally established for the SAR of the *Drosophila* histone–gene cluster, using total H1 from rat liver and DNA fragments from pBR322 as competitors (9). The H1 complement from rat liver is dominated by the subtypes H1e and H1c, while H1a, b, d and H1° are either present in low amounts or absent. In order to determine whether the six mammalian somatic subtypes, H1a–e and H1° and the more divergent male germline-specific subtype, H1t, are all SAR-binding proteins, we performed binding experiments with the minimal 657 bp SAR (75% AT) derived from the *Drosophila* histone–gene cluster and each of the purified subtypes. An HaeIII/HaeIII fragment of similar length from pUC19 (587 bp, 56% AT) was used as a competitor. The experiments consisted of mixing approximately equal amounts of the SAR and the pUC19 fragments with a limited amount of one of the subtypes. The experiments were performed at physiological salt concentration (0.14 M NaCl), where at sub-saturating protein concentrations fast-sedimenting fully complexed DNA molecules co-exist with free DNA (54). Analysis of the DNA in the soluble and insoluble fractions showed that under conditions of limited protein, H1 bound first to the SAR whilst the pUC19 fragment remained in the supernatant (Figure 1, lanes 2–15). The pUC19 appeared in the pellet only when the SAR was saturated, and was thus absent

Figure 1. Preferential binding of histone H1 subtypes to the *Drosophila* SAR. The subtypes H1a–e, H1° and H1t (lanes 2–15) show preferential binding to the SAR fragment compared with the pUC19 fragment. Once the SAR fragment is saturated, H1 binds to pUC19 (lanes 16 and 17). The subtype is indicated above. The protein/DNA ratio (w/w) is indicated. I, input mixture of the SAR and pUC19 fragments; P, pellet; S, supernatant.
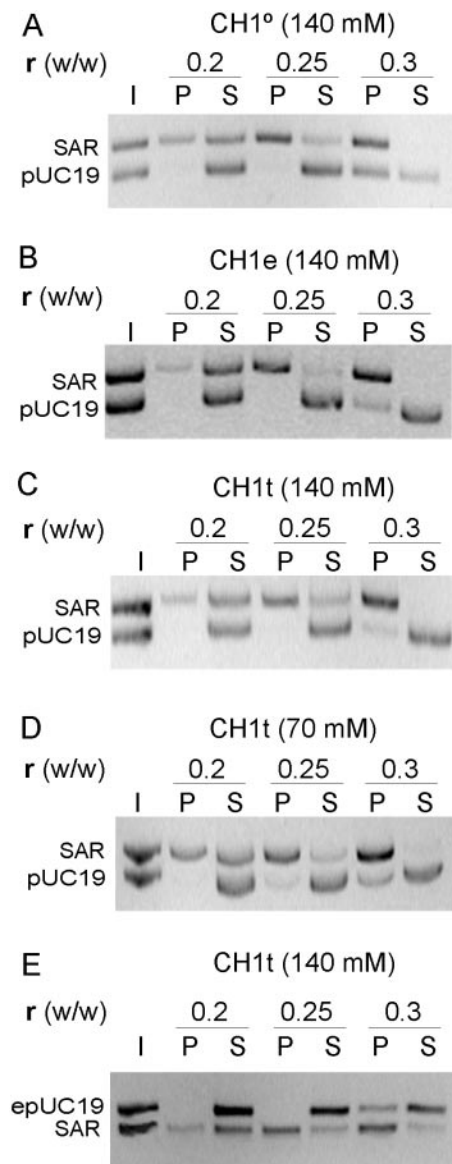

Figure 2. Preferential binding of the histone H1 C-terminal domain to the SAR fragment. (**A**) C-terminal domain of H1° (CH1°); (**B**) C-terminal domain of H1e (CH1e); (**C**, **D** and **E**) C-terminal domain of H1t (CH1t). DNA fragments are indicated: SAR (SAR fragment of 657 bp), pUC19 (pUC19 fragment of 587 bp), epUC19 (extended pUC19 fragment of 763 bp). The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the fragments; P, pellet; S, supernatant.

from the supernatants, which occurred, starting with a 1:1 mixture of SAR and pUC19, at a protein/(SAR+pUC19) weight ratio >0.5 (Figure 1, lanes 16 and 17). Above a weight ratio of 1.0, all DNA was saturated and further added protein was found as free protein in the supernatants. As shown in Figure 1, all subtypes, including H1° and H1t, showed strong cooperative binding to the SAR fragment.

The preferential binding of histone H1 is determined by its C-terminal domain

To establish the contribution of the C-terminal domain to the SAR-binding character of the entire H1 molecule, we performed experiments similar to those described for the entire H1 molecule, using the purified recombinant C-terminal domains of subtypes H1e, H1° and H1t in physiological

salt. As with the entire molecule, in competition experiments the C-terminal domains bound first to the SAR fragment, and only when it was saturated did they bind to the pUC19 fragment, which occurred, starting with a 1:1 mixture of SAR and pUC19, at a protein/(SAR+pUC19) weight ratio of about 0.3 (Figure 2). The preferential binding of the C-terminal domain to the SAR was maintained in 70 mM NaCl, although under these conditions the binding was slightly less cooperative, as indicated by the presence of a faint pUC19 band in the pellet, co-existing with a similarly faint band of SAR in the supernatant (Figure 2D). In spite of the large sequence divergence between the C-terminal domains of H1e, H1° and H1t, all showed the same strong preference for the SAR sequences (Figure 3).

To make sure that the slightly longer size of the SAR fragment over the pUC19 fragment had no influence on the

A

```
H1a  KKA---ESKAITTKVSVKAKA---SGA-AKKPKKTAGAAA-KKTV-KTPK
H1b  KKAASGEAKPKA-KKTGAAKAKKPAGATPKKPKKTAGA---KKTVKKTPK
H1c  KKAASGEAKPQA-KKAGAAKAKKPAGA-AKKPKKATGAATPKKAAKKTPK
H1d  KKAASGEAKPKA-KKAGAAKAKKPAGA-AKKPKKATGAATPKKTAKKTPK
H1e  KKAASGEAKPKA-KRAGAAKAKKPAGA-AKKPKKAAGTATAKKSTKKTPK
H1°  KGDEPKRSVAF--KKTKKEVKKVATPKKAAKPKKAASKAPSKKPKATPVK
H1t  KKAASGNDKGKG-KKSASAKAKK-MGLP-----RASR---SPKSSK--TK
     *                   *                         *           *


H1a  KPKK-PAV---SKKTSKSPKKPKVV-KAKKVAKSPAKAKAVKPKASKAKV
H1b  KAKK-PAAAG-VKKVAKSPKKAKAAAKPKKAAKSPAKPKAVKSKASKPKV
H1c  KAKK-PAAAAVTKKVAKSPKKAKVT-KPKKV-KSAS--KAVKPKAAKPKV
H1d  KAKK-PAAAAGAKKVSKSPKKVKAA-KPKKAAKSPAKAKAPKAKASKPKA
H1e  KAKK-PAAAAGAKK-AKSPKKAKAT-KAKKAPKSPAKAKTVKPKAAKPKT
H1°  KAKKKPAATP--KK-AKKPKVVKV--KPVKASKPK-KAKTVKPKAKSSAK
H1t  AVKK-PKATP--TKASGSRRKTK-GAKGVQQRKSPAKARAANP---NSGK
        **  *         *            *      *       *


H1a  TKPKTPAKP-----KKAAPKKK
H1b  TKPKT-AKPKAAKAKKAVSKKK
H1c  AKAK----------KVAAKKK
H1d  SKPKAT------KAKKAAPRKK
H1e  SKPKA-AKP-----KKTAAKKK
H1°  R--------------ASKKK
H1t  AKMVMQKTDL-----RKAAGRK
                        *
```

B

	H1b	H1c	H1d	H1e	H1°	H1t
H1a	63.1	58.2	61.5	63.1	40.2	32.0
H1b		67.2	74.6	72.1	36.9	34.4
H1c			77.0	77.0	45.9	36.1
H1d				78.7	41.0	36.1
H1e					44.3	37.7
H1°						27.9

Figure 3. Sequence identity between the C-terminal domains of mouse histone H1 subtypes. (**A**) Alignment of the C-terminal domain sequences of H1a–e, H1° and H1t. Multiple alignment was performed using the ClustalW resource under DNASTAR. The conserved positions are indicated by an asterisk (*). (**B**) Percentage similarity table.
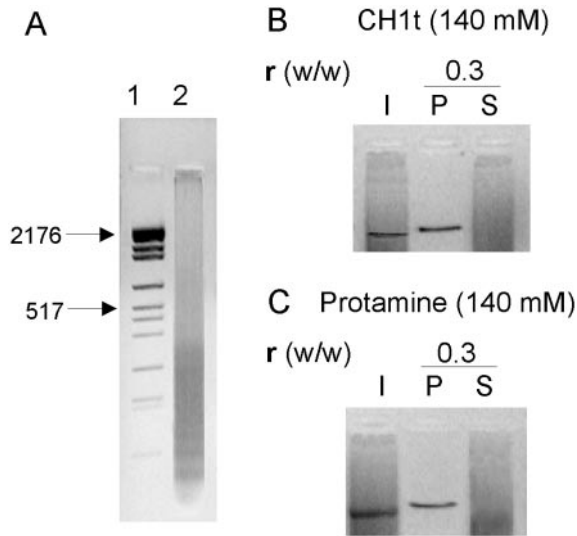
Figure 4. Preferential binding of the histone H1 C-terminal domain and protamine to the SAR in the presence of AT-rich DNA. The AT-rich DNA was obtained by polymerization of 50 bp sequence containing 75% AT and lacking A-tracts as described in Materials and Methods. (**A**) Analysis of the ligated DNA by 2% agarose gel electrophoresis; 1, DNA molecular size markers; 2, ligation reaction. The limits of the size distribution of oligomers used in the binding experiments were from 500 to 5000 bp. (**B**) Preferential binding of the C-terminal domain of H1t to the SAR. (**C**) Preferential binding of protamine to the SAR. The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the DNA fragments, the discrete band corresponds to the SAR and the smear to the AT-rich fragments; P, pellet; S, supernatant.

Figure 5. Binding of the N-terminal peptides and the globular domain of histone H1 to the SAR and pUC19 fragments. (**A**) and (**B**) N-terminal peptide of H1e (NE-1); (**C**) and (**D**) N-terminal peptide of H1° (NH-1); (**E**) globular domain of H1° (GH1°); and (**F**) globular domain of H5 (GH5). The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the SAR and pUC19 fragments; P, pellet; S, supernatant.

preference of the C-terminal domain for the SAR, we performed a competition experiment with an extended pUC19 fragment of 763 bp that included the 587 bp pUC19 sequence. It can be seen in Figure 2E that the preference of the H1 C-terminal domain for the SAR remained unaffected.

It has been reported that the preferential binding of histone H1 to SARs is not determined by its high AT content, but by the presence of abundant A-tracts (21,35). As a control of the preferential binding of the H1 C-terminal domain, we have included a competition between the SAR and a mixture of multimers of a 50 bp sequence containing 75% AT, but lacking A-tracts (see Materials and Methods). The size distribution of the multimers spanned from ∼500 to 5000 bp (Figure 4A). As shown in Figure 4B, the C-terminal domain bound with extremely high preference to the SAR in spite of the excess and longer size of a large fraction of the polymerized sequence.

In the study of the interaction of the N-terminal domain with the SAR, we used peptides that had been characterized previously (44,45). One peptide corresponded to the N-terminal domain of H1° (residues 1–20). Another peptide comprised the basic region of the N-terminal domain of H1e (residues 15–36). The affinity of these peptides for the DNA was much lower than that of the C-terminal domain, as expected from the lower number of positive charges involved in the interaction: 6–9 charges in the N-terminal domain compared with 33–46 in the C-terminal domain, depending on the subtype. Both N-terminal peptides precipitated the DNA, but a large excess of peptide was necessary. The N-terminal peptide of H1e had a moderate preference for the SAR in physiological salt (Figure 5A). The preference was most apparent at protein/
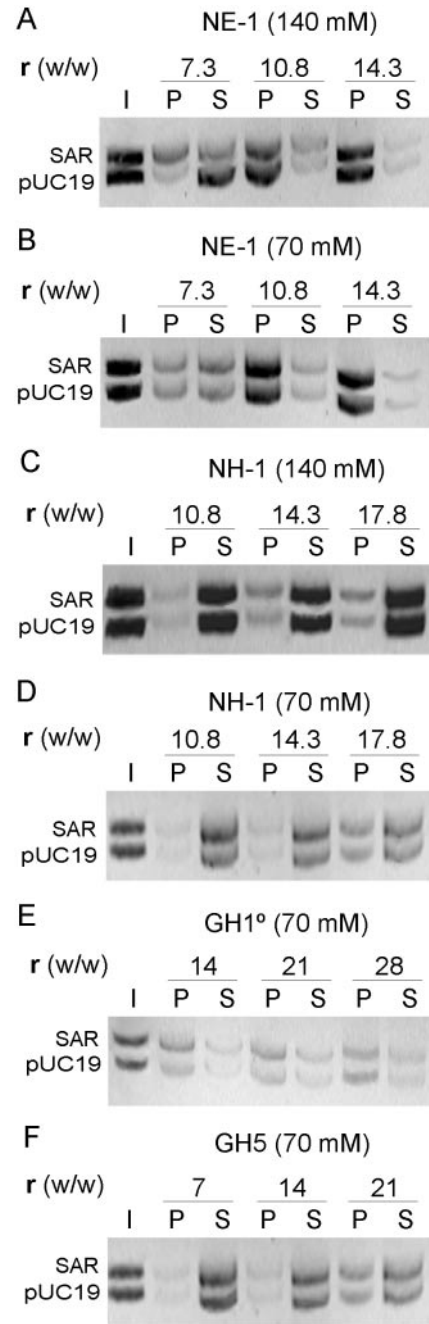
DNA ratios that precipitated a small amount of DNA, and was lost at higher protein/DNA ratios that precipitated most of the DNA. In 70 mM NaCl, the preference for the SAR fragment was completely lost (Figure 5B). The N-terminal peptide of H1° did not show any preference for the SAR, either in 140 mM or in 70 mM NaCl (Figure 5C and D).

With the globular domain, precipitation experiments could not be performed in 140 mM NaCl because the interaction
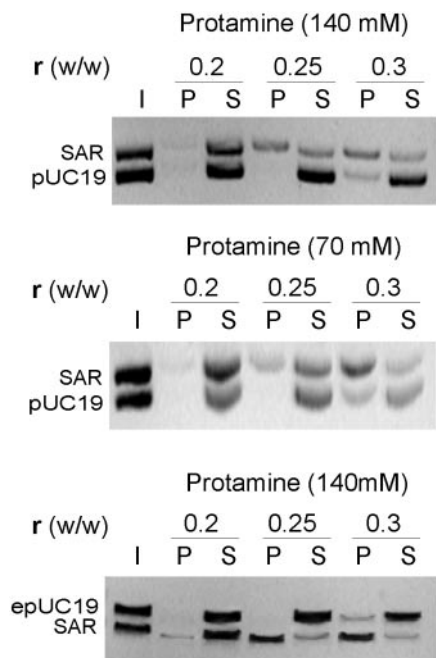
Figure 6. Preferential binding of protamine to the SAR fragment. DNA fragments are indicated: SAR (SAR fragment of 657 bp), pUC19 (pUC19 fragment of 587 bp), epUC19 (extended pUC19 fragment of 763 bp). The concentration of NaCl is indicated in parenthesis. The protein/DNA ratio (w/w) is indicated. I, input mixture of the fragments; P, pellet; S, supernatant.

with the DNA was too weak in this salt concentration. Binding experiments were carried out instead in 70 mM NaCl, where with a high enough protein/DNA ratio all the DNA could be precipitated. The globular domain of H1° showed no preference for the SAR fragment, not even at protein/DNA ratios that precipitated a small proportion of the DNA (Figure 5E). The globular domain of H5, an avian subtype closely related to H1°, also failed to bind preferentially to the SAR (Figure 5F).

Protamines also bind preferentially to SARs

Typical protamines are sperm basic proteins with an arginine content of 60% or higher. We used salmine in the binding experiments, a typical protamine from salmon sperm. It has 32 residues, 21 of which are arginine. The Arg residues are mostly present in clusters of up to six residues. Figure 6 shows that protamine also bound with high preference to the SAR fragment either in physiological salt or in 70 mM NaCl. The preference is, however, not so high as displayed by the C-terminal domain, as shown by the presence of a small amount of competitor pUC19 in the pellets coexisting with a small amount of SAR in the supernatants. The preference for the SAR was maintained when the longer pUC19 fragment (763 bp) was used in the experiments (Figure 6). As in the case of the H1 C-terminal domain, protamine also bound preferentially to the polymerized sequence lacking A-tracts (Figure 4C).

Distamycin abolishes the SAR-binding character of the C-terminal domain of H1 and protamine

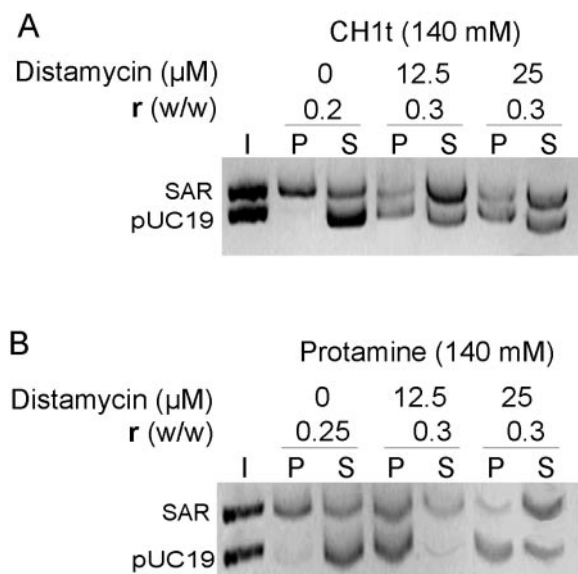Distamycin is an antibiotic drug with high selectivity for A-tract DNA (55,56). Binding to A-tracts by distamycin

Figure 7. Distamycin inhibits the preferential binding of the histone H1 C-terminal domain and protamine to the SAR fragment. (**A**) C-terminal domain of H1t (CH1t); (**B**) salmon protamine. The concentration of NaCl is indicated in parenthesis. The concentration of distamycin and the protein/DNA ratio (w/w) are indicated. I, input mixture of the SAR and pUC19 fragments; P, pellet; S, supernatant.

abolishes the specific interaction of SARs with the nuclear scaffold and SAR binding proteins. In particular, binding of distamycin results in the suppression of preferential binding of histone H1 to SAR containing DNA, leading to a redistribution of histone H1 between SAR and non-SAR DNA (31). Distamycin has been used to confirm the SAR-binding properties of several proteins. We examined the effect of distamycin on the binding of the C-terminal domain of H1 and protamine, which were shown to bind highly preferentially to the SAR fragment. In both cases, distamycin abolished the preferential binding of the proteins to the SAR (Figure 7). The effect was so clear as to reverse the binding preferences, binding to pUC19 being stronger than to the SAR in the presence of distamycin.

DISCUSSION

Histone H1 has been generically described as an SAR-binding protein. Here, we have shown that the highly conserved mammalian subtypes, H1a–e, and the highly divergent H1° and H1t are all SAR-binding proteins. The C-terminal domain of histone H1 contributes to a large extent to the binding free energy of the entire molecule to DNA, providing the localized charge neutralization of the DNA necessary for the chromatin folded state (57–59). It was thus of interest to investigate whether the SAR-binding character of H1 subtypes was a property of the C-terminal domain. We studied the C-terminal domains of subtypes H1e, the most abundant somatic subtype, H1°, a subtype that accumulates in quiescent cells, and H1t, the male germline-specific subtype. In spite of their large sequence divergence (44.3% sequence identity between H1e and H1°, and 27.9% between H1° and H1t; Figure 3), all three C-terminal domains appear to have very strong preference for

the SAR. H1e and H1° contain SPKK motifs, presumably contributing to SAR-binding specificity; however, this motif is not present in H1t, which indicates that its presence is not essential to determine the SAR-binding character of histone H1 subtypes. As shown previously for the entire molecule (31), distamycin, a minor groove binder with high affinity for A-tracts, abolishes the preferential binding of the C-terminal domain to the SAR fragment. This suggests involvement of the DNA minor groove in the interaction with the C-terminal domain, although it does not exclude the possibility that the C-terminal domain also binds to the wide groove (60). As shown by X-ray crystallography, distamycin occupies the minor groove with preference for narrow minor grooves (32). However, distamycin may also cause subtle changes in the structural parameters and mechanoelastic properties of the DNA that could contribute to the suppression of the preferential binding of SAR-binding proteins (56).

In agreement with its reduced number of positive charges compared with the C-terminal domain, the N-terminal domain exhibits much weaker DNA binding, and a large excess of protein is necessary to precipitate the DNA. The N-terminal peptide of H1° showed no preference for the SAR fragment. The N-terminal domain of H1e had a moderate preference for the SAR fragment in physiological salt, but in 70 mM salt the binding preference was completely lost. With the globular domain, binding experiments had to be performed at 70 mM NaCl, as in 140 mM the binding was too weak to observe a precipitate even with a large excess of protein. Under these conditions, the globular domain did not show any preference for the SAR sequences over the pUC19 fragments. Given the low selectivity and weak binding of the N-terminal peptides and the globular domain, it is unlikely that they make a substantial contribution to the SAR-binding character of the whole protein. It thus appears that the preferential binding of histone H1 to SARs is basically determined by its C-terminal domain.

Phosphorylation is the main post-translational modification undergone by histone H1. Most phosphorylation sites are located in the C-terminal domain. Phosphorylation weakens the binding of H1 to DNA, and could thus facilitate SAR chromatin opening through cooperative H1 dissociation (61,62). The C-terminal domain could also be responsible for the targeting of H1 molecules to SARs. At substoichiometric concentrations, the higher affinity of H1 for SARs would guarantee the saturation of SAR sequences with H1 in preference to other sequences.

Evolutionary evidence has recently been obtained for chordate protamines having originated from histone H1 through general substitution of Lys residues by Arg (46). In view of the preferential binding of the C-terminal domain of H1 to SARs, we investigated whether such a property was conserved in protamines. The results show that protamine also binds to the SAR fragment with high specificity. Moreover, the binding preference is also determined by the oligo-A tracts as shown by the suppression of preferential binding to the SAR by distamycin. The effect is so strong as to reverse the binding preferences of protamine. This is indicative of minor groove recognition by protamine. The role, if any, of the preferential binding of protamines to SARs cannot be ascertained at this stage; however, it is possible that this property could lend some spatial and temporal order to the substitution of core

histones by protamines in spermatogenesis, the SARs providing nucleation sites for the substitution process.

The features that confer SAR-binding specificity to a protein or protein motif have yet to be established. Mammalian HMGA non-histone proteins contain multiple copies of a DNA-binding motif called 'AT hook' that preferentially binds to the narrow minor groove of AT sequences (63); the core of the AT hook is the sequence Arg–Gly–Arg. The interaction of the RGR element with the minor groove has been characterized by NMR spectroscopy (64). No RGA motifs are found in either histone H1 or protamine. While histone H1 is lysine-rich, the core motif of AT hooks and protamine contain Arg. The C-terminal domain is an even better SAR-binding protein than protamine. The choice between Arg and Lys does not therefore appear essential in itself in determining the SAR-binding character. Presumably, the interaction should as a rule involve minor groove recognition, as all SAR-binding proteins so far described are competed by distamycin. The narrower minor groove associated with A-tracts would be preferred because it would give better van der Waals contacts between the walls of the minor groove and the ligands (32). In the case of highly cationic ligands, such as the C-terminal domain of histone H1 and protamine, potentially interacting with the DNA phosphates, a narrower minor groove would also offer a more intense electrostatic potential along the minor groove path that could also contribute to the preferential binding.

REFERENCES

1. Zlatanova,J. and van Holde,K. (1992) Histone H1 and transcription: still an enigma? *J. Cell Sci*., **103**, 889–895.
2. Bouvet,L., Dimitrov,S. and Wolffe,A.P.K. (1994) Specific regulation of Xenopus chromosomal 5S rRNA gene transcription *in vivo* by histone H1. *Genes Dev*., **8**, 1147–1159.
3. Khochbin,S. and Wolffe,A.P. (1994) Developmentally regulated expression of linker-histone variants in vertebrates. *Eur. J. Biochem*., **225**, 501–510.
4. Shen,X. and Gorovsky,M.A. (1996) Linker histone H1 regulates specific gene expression but not global transcription *in vivo*. *Cell*, **86**, 475–483.
5. Wolffe,A.P., Khochbin,P.S. and Dimitrov,S. (1997) What do linker histones do in chromatin? *BioEssays*, **19**, 249–255.
6. Lee,H.L. and Archer,T.K. (1998) Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. *EMBO J*., **17**, 1454–1466.
7. Vermaak,D., Steinbach,O.C., Dimitrov,S., Rupp,R.A.W. and Wolffe,A.P. (1998) The globular domain of histone H1 is sufficient to direct specific gene repression in early *Xenopus* embryos. *Curr. Biol*., **8**, 533–536.
8. Dou,Y., Mizzen,C.A., Abrams,M., Allis,C.D. and Gorovsky,M.A. (1999) Phosphorylation of linker histone H1 regulates gene expresión *in vivo* by mimicking H1 removal. *Cell*, **4**, 641–647.
9. Izaurralde,E., Käs,E. and Laemmli,U.K. (1989) Highly preferential nucleation of histone H1 assembly on scaffold-associated regions. *J. Mol. Biol*., **210**, 573–585.

10. Pennings,S., Meersseman,G. and Bradbury,E.M. (1994) Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proc. Natl Acad. Sci. USA*, **91**, 10275–10279.

11. Panyim,S. and Chalkley,R. (1969) A new histone found only in mammalian tissues with little cell division. *Biochem. Biophys. Res. Commun.*, **37**, 1042–1049.

12. Bucci,L.R., Brock,W.A. and Meistrich,M.L. (1982) Distribution and synthesis of histone 1 subfractions during spermatogenesis in the rat. *Exp. Cell Res.*, **140**, 111–118.

13. Lennox,R.W. (1984) Differences in evolutionary stability among mammalian H1 subtypes. Implications for the roles of H1 subtypes in chromatin. *J. Biol. Chem.*, **259**, 669–672.

14. Tanaka,M., Hennebold,J.D., Macfarlane,J. and Adashi,E.Y. (2001) A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (CS-H1) of sea urchin and the B4/H1M histone of the frog. *Development*, **128**, 655–664.

15. Khochbin S. (2001) Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene*, **27**, 11–12.

16. Lennox,R.W., Oshima,R.G. and Cohen,L.H. (1982) The H1 histones and their interphase phosphorylated states in differentiated and undifferentiated cell lines derived from murine teratocarcinomas. *J. Biol. Chem.*, **257**, 5183–5189.

17. Hall,J.M. and Cole,R.D. (1985) Modulation in proportions of H1 histone subfractions by differential changes in synthesis and turnover during butyrate treatment of neuroblastoma cells. *Biochemistry*, **24**, 7765–7771.

18. Domínguez,V., Piña,B. and Suau,P. (1992) Histone H1 subtype synthesis in neurons and neuroblasts. *Development*, **115**, 181–185.

19. Ponte,I., Vidal-Taboada,J.M. and Suau,P. (1998) Evolution of the H1 histone class: evidence for the functional differentiation of the subtypes. *Mol. Biol. Evol.*, **15**, 702–708.

20. Liao,L.W. and Cole,R.D. (1981) Differences among fractions of H1 histones in their interactions with linear and superhelical DNA: circular dichroism. *J. Biol. Chem.*, **256**, 10124–10128.

21. Khadake,J.R., Manchanahalli,R. and Rao,M.R.S. (1995) DNA- and chromatin-condensing properties of rat testes H1a and H1t compared to those of rat liver H1bdec; H1t is a poor condenser of chromatin. *Biochemistry*, **34**, 15792–15801.

22. Talasz,H., Sapojnikova,N., Helliger,W., Linder,H. and Puschendorf,B. (1998) *In vitro* binding of H1 histone subtypes to nucleosomal organized mouse mammary tumor virus long terminal repeat promotor. *J. Biol. Chem.*, **273**, 32236–32243.

23. Mirkovitch,J., Mirault,M.-E. and Laemmli,U.K. (1984) Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell*, **39**, 223–232.

24. Strick,R. and Laemmli,U.K. (1995) SARs are cis DNA elements of chromosome dynamics: synthesis of a SAR repressor protein. *Cell*, **83**, 1137–1148.

25. Adachi,Y., Käs,E. and Laemmli,U.K. (1989) A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J.*, **8**, 3997–4006.

26. Luderus,M.E., de Graaf,A., Mattia E., den Blaauwen J.L., Grande M.A., de Jong,L. and van Driel,R. (1992) Binding of matrix attachment regions to lamin B1. *Cell*, **70**, 949–959.

27. Dickinson,L.A. and Kohwi-Shigematsu,T. (1995) Nucleolin is a matrix attachment region DNA-binding protein that specifically recognizes a region with high base-unpairing potential. *Mol. Cell. Biol.*, **15**, 456–465.

28. Zhao,K., Kas,E., Gonzalez,E. and Laemmli,U.K. (1993) SAR-dependent mobilization of histone H1 by HMG-I/Y *in vitro*: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J.*, **12**, 3237–3247.

29. Dickinson,L.A., Joh,T., Kohwi,Y. and Kohwi-Shigematsu,T. (1992) A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell*, **70**, 631–645.

30. Kipp,M., Gohring,F., Ostendorp,T., van Drunen,C.M.van Driel,R., Przybylski,M. and Fackelmayer,F.O. (2000) SAF-Box, a conserved protein domain that specifically recognizes scaffold attachment region DNA. *Mol. Cell. Biol.*, **20**, 7480–7489.

31. Käs,E., Izaurralde,E. and Laemmli,U.K. (1989) Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin. The role of oligo(dA).oligo(dT) tracts. *J. Mol. Biol.*, **210**, 587–599.

32. Coll,M., Frederick,C.A., Wang,A.H.J. and Rich,A. (1987) A bifurcated hydrogen-bonded conformation in the (dA.dT) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl Acad. Sci. USA*, **84**, 8385–8339.

33. Nelson,H.C., Finch,J.T., Luisi,B.F. and Klug,A. (1987) The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature*, **330**, 221–226.

34. Varga-Weisz,P., van Holde,K. and Zlatanova,J. (1993) Preferential binding of histone H1 to four-way helical junction DNA. *J. Biol. Chem.*, **268**, 20699–20700.

35. Yaneva,J., Schroth,G.P., van Holde,K.E. and Zlatanova,J. (1995) High-affinity binding sites for histone H1 in plasmid DNA. *Proc. Natl Acad. Sci. USA.*, **92**, 7060–7064.

36. Ivanchenko,M., Zlatanova,J. and van Holde,K.E. (1997) Histone H1 preferentially binds to superhelical DNA molecules of higher compaction. *Biophys. J.*, **72**, 1388–1395.

37. Hartman,P.G., Chapman,G.E., Moss,T. and Bradbury,E.M. (1977) Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The three structural regions of the histone H1 molecule. *Eur. J. Biochem.*, **77**, 45–51.

38. Subirana,J.A. (1990) Analysis of the charge distribution in the C-terminal region of histone H1 as related to its interaction with DNA. *Biopolymers*, **29**, 1351–1357.

39. Ramakrishnan,V., Finch,J.T., Graziano,V., Lee,P.L. and Sweet,R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219–223.

40. Clore,G.M., Gronenborn,A.M., Nilges,M., Sukumaran D.K. and Zarbock,J. (1987) The polypeptide fold of the globular domain of histone H5 in solution. A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO J.*, **6**, 1833–1842.

41. Cerf,C., Lippens,G., Muyldermans,S., Segers,A., Ramakrishnan,V., Wodak,S.J., Hallenga,K. and Wyns,L. (1993) Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: sequential assignment and secondary structure. *Biochemistry*, **32**, 11345–11351.

42. Vila,R., Ponte,I., Jiménez,M.A., Rico,M. and Suau,P. (2000) A helix-turn motif in the C-terminal domain of histone H1. *Protein Sci.*, **9**, 1–10.

43. Vila,R., Ponte,I., Collado,M., Arrondo,J.L.R. and Suau,P. (2001) Induction of secondary structure in a COOH-terminal peptide of histone H1 by interaction with the DNA. *J. Biol. Chem.*, **276**, 30898–30903.

44. Vila,R., Ponte,I., Collado,M., Arrondo,J.L.R., Jiménez,M.A., Rico,M. and Suau,P. (2001) DNA-induced α-helical structure in the N-terminal domain of histone H1. *J. Biol. Chem.*, **276**, 46429–46435.

45. Vila,R., Ponte,I., Jiménez,M.A., Rico,M. and Suau,P. (2001) An inducible Helix-Gly-Gly-Helix motif in the N-terminal domain of histone H1e: a CD and NMR study. *Protein Sci.*, **11**, 214–220.

46. Lewis,J.D., Saperas,N., Song,Y., Zamora,M.J., Chiva M. and Ausió,J. (2004) Histone H1 and the origin of protamines. *Proc. Natl Acad. Sci. USA*, **101**, 4148–4152.

47. Thompson,R.J. (1973) Studies on RNA synthesis in two populations of nuclei from the mammalian cerebral cortex. *J. Neurochem.*, **21**, 19–40.

48. Piña,B., Martínez,P., Simón,L. and Suau,P. (1984) Differential kinetics of histone H1° accumulation in neuronal and glial cells from rat cerebral cortex during postnatal development. *Biochem. Biophys. Res. Commun.*, **123**, 679–702.

49. García-Ramírez,M., Leuba S.H. and Ausio J. (1990) One-step fractionation method for isolating H1 histones from chromatin under nondenaturing conditions. *Protein Expr. Purif.*, **1**, 40–44.

50. Böhm,E.L., Strickland,W.N., Strickland,M., Thwaits,B., Van der Westhuizen,D.R. and Von Holt,C. (1973) Purification of the five main calf thymus histone fractions by gel exclusion chromatography. *FEBS Lett.*, **34**, 217–221.

51. Brown,D.T., Alexander,B.T. and Sittman,D.B. (1996) Differential effect of H1 variant overexpression on cell cycle progression and gene expression. *Nucleic Acids Res.*, **24**, 486–493.

52. Piña,B.Martínez,P. and Suau,P. (1987) Changes in H1 complement in differentiating rat brain cortical neurons. *Eur. J. Biochem.*, **164**, 71–76.

53. Goytisolo,F.A., Gerchman,S.E., Yu,X., Rees,C., Graciano,V., Ramakrishnan,V. and Thomas,J.O. (1996) Identification of two DNA binding sites on the globular domain of histone H5. *EMBO J.*, **15**, 3421–3429.

54. Rodríguez,A.T., Pérez,L., Morán,F., Montero,F. and Suau,P. (1991) Cooperative interaction of the C-terminal domain of histone H1 with DNA. *Biophys. Chem.*, **39**, 145–152.

55. Van Dyke,M.W., Hertzberg,R.P. and Dervan,P.B. (1982) Map of distamycin, neotropsin and actinomycin binding site on heterogeneous DNA: DNA clevage-inhibition patterns with methidiumpropyl-EDTA.Fe(II). *Proc. Natl Acad. Sci. USA*, **79**, 5470–5474.

56. Fox,K.R. and Waring,M.J. (1984) Kinetic evidence for redistribution of actinomycin molecule between potential binding sites. *Nucleic Acids Res.*, **12**, 9271–9285.

57. Lu,X. and Hansen,J.C. (2003) Revisiting the structure and functions of the linker histone C-terminal tail domain. *Biochem. Cell Biol.*, **81**, 173–176.

58. Lu,X. and Hansen,J.C. (2004) Identification of specific functional subdomains within the linker histone H1° C-terminal domain. *J. Biol Chem.*, **279**, 8701–8707.

59. Hendzel,M.J., Lever,M.A., Crawford,E. and Th'ng,J.P. (2004) The C-terminal domain is the primary determinant of histone H1 binding to chromatin *in vivo*. *J. Biol. Chem.*, **279**, 20028–20034.

60. Mamoon,N.M., Song,Y. and Wellman,S.E. (2002) Histone H1(0) and its carboxyl-terminal domain bind in the major groove of DNA. *Biochemistry*, **41**, 9222–9228.

61. Hill,C.S., Rimmer,J.M., Green,B.N., Finch,J.T. and Thomas,J.O. (1991) Histone-DNA interactions and their modulation by phosphorylation of –Ser-Pro-X-Lys/Arg. *EMBO J.*, **10**, 1939–1948.

62. Laemmli,U.K., Käs,E., Poljak,L. and Adachi,Y. (1992) Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Curr. Opin. Genet. Dev.*, **2**, 275–285.

63. Reeves,R. and Nissen,M.S. (1990) The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J. Biol. Chem.*, **265**, 8573–8582.

64. Huth,J.R., Bewley,C.A., Nissen,M.S., Evans,J.N., Reeves,R., Gronenborn,A.M. and Clore,G.M. (1997) The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nature Struct. Biol.*, **8**, 657–665.