***In vivo* protein binding sites and nuclease hypersensitivity in the promoter region of a cell cycle regulated human H3 histone gene**

Urs Pauli⁺, Susan Chrysogelos, Harry Nick, Gary Stein* and Janet Stein*

University of Florida, College of Medicine, Gainesville, FL 32610, USA

Received December 23, 1987; Revised and Accepted February 15, 1989

ABSTRACT

The chromatin structure and protein-DNA interactions of a cell cycle regulated human H3 histone gene have been examined at different levels of resolution. Using traditional Southern blot analysis we have investigated the accessibility of the H3 coding region and its flanking sequences to DNase I, S1 nuclease and restriction endonuclease digestion. Using the native genomic blotting method recently developed in our laboratory, two sites of protein-DNA interaction in the proximal 240 bp of the promoter region of this H3 gene were established. Further *in vivo* analysis of protein-DNA binding sites in intact cells by genomic sequencing revealed, with single nucleotide resolution, the guanine contacts and footprints of the proteins bound to the promoter. The relative locations of protein-DNA interactions in this H3 gene are similar to those identified *in vivo* and *in vitro* in a cell cycle dependent human H4 histone gene. The proteins complexed with the H3 histone gene promoter can be dissociated between 0.16 and 0.28 M NaCl. The protein-DNA contacts persist throughout the cell cycle and thus may have a functional relationship with the basal level of transcription of this H3 gene that occurs during and outside of S phase.

INTRODUCTION

There is a temporal and functional coupling of the expression of cell cycle dependent human histone genes with DNA replication, with control mediated at both transcriptional and post-transcriptional levels (1-4; for review see 5). These cell cycle dependent genes are transcribed throughout the cell cycle at a basal level, but exhibit a 3-5 fold increase in transcription during the initial part of S phase (1,2,4). Post-transcriptional control of histone gene expression is reflected by the 20-100 fold increase in cellular histone mRNA levels during S phase, and by the rapid and selective destabilization of histone mRNA that occurs at the natural termination of S phase or following inhibition of DNA synthesis (6). An understanding of histone gene regulation, therefore, requires the identification of chromatin complexes and DNA sequences that influence levels of transcription, and of the mechanisms by which histone mRNA stability is coupled to DNA synthesis.

Transcriptional activity of genes can be correlated to a series of modifications in chromatin structure such as DNase I (hyper)sensitivity (7-9), disruption of the nucleosomal pattern (10), S1 nuclease sensitivity (11), restriction enzyme accessibility (12) and protein-DNA binding sites (13). These modifications are associated with actively expressed genes, potentially active genes, and genes that were once active (for review see (14)). Such structural changes are involved in the active transcription of a gene at different levels: i) changes that open the overall chromatin structure to permit transcription (8), ii) changes that alter chromatin structure in the immediate promoter regions to control the expression of a gene, and iii) changes that allow positive and negative regulatory elements to modulate transcription at a distance from the proximal promoter (15,16).

Transcriptional regulation has been shown in several systems to involve a series of cis-acting elements interacting with sequence-specific binding proteins (for review see (17)). For human histone genes such elements have been identified and DNA binding proteins have been partially purified (18-21). To investigate the changes in chromatin structure that may be involved in the regulation of histone genes, we examined the sites of nuclease sensitivity and protein-DNA interaction of a cell cycle dependent human H3 histone gene designated ST519. Dimethylsulfate (DMS) (22), DNase I (7,23), S1 nuclease (11) and restriction endonuclease (12) analyses were carried out in combination with Southern blotting (24), native genomic blotting (25,26) and genomic sequencing (22). This combined approach allowed us to screen the coding and flanking regions of this H3 histone gene over several kilobases for (hyper)sensitive and protected regions, and then to analyze areas of interest for protein-DNA interactions at intermediate or single nucleotide resolution.

Our results indicate the presence within the promoter region of this H3 histone gene of one strong DNase I hypersensitive site containing two, and possibly three, distinct sites of protein-DNA interaction with different properties. We have established the contact points of these proteins with specific sequences, as well as the regions that are protected by these proteins from DNase I digestion. The specific protein-DNA complexes persist throughout the cell cycle and can be dissociated by treatment of nuclei with sodium chloride concentrations exceeding 0.2 M. Our results suggest that the protein binding sites found in close proximity to the mRNA initiation site are functionally related to the basal level of transcription of this ST519 H3 histone gene. The binding sites in this H3 gene are similar to those in a cell cycle regulated human H4 histone gene, although only the more upstream sites

contain any sequence similarities -- the sites are located in similar regions of the promoter, are dissociated from their binding factors at the same salt concentrations, and contain dyad symmetries in the DNase I footprints.

MATERIALS AND METHODS

Isolation and treatment of nuclei

HeLa S3 cells were synchronized by a double thymidine block procedure (27) and harvested at various times after release from the second block. Isolation of nuclei (28), extraction of nuclei with various salt concentrations (29), nuclease digestions (28,30) and purification of the DNA (28) were done as previously described. Experiments included appropriate "no enzyme" or deproteinized DNA controls.

Native genomic blotting and probe preparation

Restricted genomic DNA (10-15 μ g) was separated in a 4% native polyacrylamide gel in 50 mM TBE buffer (50 mM Tris-HCl, pH 8.3, 50 mM boric acid, 1 mM EDTA). The gel was boiled to denature the DNA and electrotransferred to a nylon membrane (Genescreen, New England Nuclear) (See ref. 26 for details). UV crosslinking, preparation of the probes and hybridization were done according to the procedure of genomic sequencing (22). The M13 probes are identified in Figure 3 and are specified in the figure legends. The EcoRI/PvuII probe for the alkaline Southern blot (31) shown in Figure 1 was an isolated fragment that was labeled by the random oligo primer extension method (32). All probes chosen for use are single copy and recognize only the ST519 flanking regions, not those of other H3 histone genes. Exposures of autoradiograms were done at -70°C with an intensifying screen.

DMS Fingerprinting and DNase I Footprinting

Treatment of the cells with DMS, separation of the DMS and DNase I treated DNAs on genomic sequencing gels, electrotransfer to nylon (Genescreen), and probe preparation and hybridization were done as described (22,23,29,33).

RESULTS

Nuclease sensitivity of the ST519 H3 histone gene and its flanking regions

The chromatin structure of the H3 histone gene ST519 was investigated for the presence of nuclease (hyper)sensitive sites. Approximately 2.5 kb of sequences surrounding the coding region (including 800 bp of 5' flanking region and 1500 bp of 3' flanking region) were screened for sensitivity to DNase I, S1 nuclease and the restriction enzyme MnlI during the cell cycle. HeLa S3 cells were synchronized at the G1/S phase boundary, and nuclei were prepared from

Figure 1. DNase I, S1 nuclease and restriction endonuclease sensitivity of the ST519 human H3 histone gene and its flanking regions. Nuclei were isolated from synchronized HeLa S3 cells at early S phase, mid S phase and mitosis/G1 phase (1, 5 and 10 hrs, respectively, after release from a double thymidine block). DNase I treatment was with 3.5, 5 or 7.5 $\mu\text{g}/\text{ml}$ for 10 min at 37°C (lanes 2-4, early S phase; lanes 5-7, mid S phase; lanes 8-10, mitosis/G1). No pattern was observed for a "no enzyme" control sample. S1 nuclease digestion was with 5000 U/ml for 30 min at 37°C (lane 11, early S; lane 12, mid S; lane 13, mitosis/G1). MnlI was added at 450 U/ml for 30 min at 30°C (lane 15, early S; lane 16, mid S; lane 17, mitosis/G1). Lane 14: Intact purified HeLa DNA digested with EcoRI and then partially digested with MnlI. Digestion of deproteinized DNA with DNase I or S1 nuclease did not produce a detectable pattern. Lane 1: EcoRI/HindIII-digested lambda DNA. DNA was isolated from nuclei, digested with EcoRI, separated in a 1.2% agarose gel, and alkaline blotted to nylon (Zetabind) (31). The filter was hybridized with the PvuII/EcoRI fragment indicated. The map shows the H3 coding and flanking regions and the positions of several MnlI sites. E = EcoRI, Pv = PvuII.

early S phase, mid S phase, and mitosis/G1 phase cells and treated with various nucleases. The purified DNA was restricted, blotted to nylon filters, and hybridized with the EcoRI/PvuII fragment indicated in Figure 1 as an indirect end label. This probe recognizes uniquely the ST519 H3 histone gene. The restriction enzyme MnlI was chosen for these studies because it has numerous recognition sites throughout the H3 coding and flanking regions, generating a ladder of bands.

There is one main nuclease hypersensitive region in the immediate promoter of this H3 histone gene, which is recognized by DNase I, S1 and MnlI (Figure

Figure 2: High resolution analysis by native genomic blotting of the DNase I hypersensitive area of the ST519 H3 histone gene. The DNA was restricted with *Dra*I and hybridized with the *Dra*I 5' upper probe (see Figure 3). Lane 1: DNA from nuclei of thymidine blocked cells digested with 7 μ g/ml DNase I. Lanes 2-7: DNA from nuclei of cells at 1 (early S phase), 3, 6 (mid S phase), 8, 10 and 12 hrs (mitosis/G1 phase) after release of the thymidine block, treated with 5-10 μ g/ml DNase I. Lane 8: Deproteinized HeLa DNA digested with *Dra*I and partially with *Hinc*II. Restriction enzymes are: D = *Dra*I, Hc = *Hinc*II. Lane 5: I, IIa and IIb refer to three light areas; the IIb region is not so well defined as the other two (see text). On the right the three areas are shown relative to the coding region of the H3 histone gene. ATG = translation start codon. H3 = coding region; only part of the coding region is shown.

1). The strongest DNase I sensitivity centers around the *Mn*I site that maps to -160 bp with respect to the ATG initiation codon. The S1 site covers a larger region from approximately -180 bp to +60 bp. Upstream of the hypersensitive region, there is strong protection except for an area around -500 bp that is accessible (same accessibility as the 3' flanking region) to both DNase I and S1 nuclease. Downstream, the coding region continues to be sensitive to all nucleases, though at a lower level than the immediate promoter region. Accessibility decreases further in the region 3' to the coding sequences, reaching minimal sensitivity downstream of the gene at +1200 bp. The pattern of the *Mn*I digest parallels the sensitivity of DNase I and S1 nuclease: whereas the site at -160 around the hypersensitive area is more sensitive in nuclei, the sites in the 3' flanking region and the site around -500 bp are less accessible in nuclei than in the deproteinized control digest.

During the cell cycle, there are no dramatic changes in the core nuclease hypersensitive site located at -160 bp upstream of the ATG codon, paralleling

the case with the FO108 H4 histone gene that has a hypersensitive site in an analogous region (28). However, the sensitivity of both coding regions to DNase I and S1 nuclease is greatest during mid S phase and lowest during mitosis/G1. Additionally, both genes show increasing protection in their 3' flanking regions, and each has an area slightly sensitive to both DNase I and S1 nuclease upstream of the hypersensitive site. In the H3 gene this site is at -500 bp, while in the H4 gene it is located farther upstream at -720 bp (30). Notably, except for this one area, the region upstream of the hypersensitive site in the H3 gene is much more protected from digestion than in the H4 gene.

Native genomic blotting displays two regions of protein-DNA interactions in the first 240 bp upstream of the ATG codon

Having detected one site of DNase I hypersensitivity in the promoter region of this ST519 H3 histone gene, we focused our experiments on this area. The native genomic blotting technique recently developed in this laboratory (25,26) was used to scan 400 bp of the promoter and coding region at intermediate resolution (Figure 2).

Nuclei from HeLa S3 cells in early S phase, mid S phase and mitosis/G1 phase were treated with DNase I, and the DNA was purified and digested with DraI prior to native genomic blotting (see Materials and Methods). When such a blot was hybridized with the DraI 5' upper probe (Figure 3), two light (protected) areas (and a possible third area) representing sites of protein-DNA interaction were visible in the autoradiogram (Figure 2) and are designated I, IIa and IIb, respectively (This nomenclature was chosen because the sequences underlying IIa and IIb both include a CCAAT element within their footprint). The pattern observed for digestion of deproteinized DNA was quite different (data not shown) and supported the designation of the light areas as sites of protein protection. The light areas in Figure 2 are located within the region of the one strong DNase I hypersensitive band in the agarose blot shown in Figure 1. Site I mapped to a region from -225 to -180 bp, site IIa from -140 to -115 and site IIb was estimated to map from -95 to -80 bp, while the main band of DNase I hypersensitivity was located between site I and site IIa. Not all protected areas displayed the same properties: site I was very distinct in all lanes, whereas sites IIa and especially IIb had much more background; in fact, on some blots site IIb was difficult to detect. This suggests that interactions of the protein(s) with sites IIa (and IIb) are weaker than interactions occurring at site I. The pattern of the light areas did not change during G1, S or mitosis, indicating that these protein-DNA interactions

Figure 3: Probes used for native genomic blotting and genomic sequencing of the H3 histone gene. Top: upper and lower strands of the cell cycle dependent H3 histone gene ST519. CAAT, TATA = consensus elements in the promoter region. AUG = translation start codon. The dark box corresponds to the transcribed region. Bottom: restriction map of important sites in this gene. D = DraI, Hc = HincII, Hd = HindIII and N = NcoI. The solid bars with the arrows display the probes used in our experiments. The end of each bar corresponds to one restriction site. The nomenclature is according to Church and Gilbert (22).

persist throughout the cell cycle. The native genomic blot also confirms the earlier observation that the sequences upstream of the hypersensitive region are resistant to DNase I digestion while the coding region remains sensitive.

To confirm that the light areas represented sites of protein-DNA interactions, we performed experiments in which the nuclei were washed with various concentrations of salt prior to DNase I digestion in order to dissociate the protein-DNA complexes (data not shown). Analysis of these results demonstrated that NaCl at concentrations higher than 0.2 M removed the light areas from the native genomic blots.

The pattern of protein-DNA interactions detected in the H3 histone gene promoter was somewhat different from the pattern we have observed with an H4 histone gene (29). Two distinct sites with similar binding properties can be seen in the H4 histone gene promoter, whereas in the H3 histone promoter the sites have different binding properties. The size and localization of the sites also seem to be different for the two genes, e.g., H4 site I= 35 bp at -150 bp, H3 site I= 45 bp at -235 bp.

Protein-DNA contacts at single nucleotide resolution on the upper and lower strands of the H3 histone gene promoter

To characterize further the protein binding sites of the H3 histone gene promoter, we used the genomic sequencing method (DMS fingerprinting, DNase I footprinting) to display protein-DNA interactions at single nucleotide

resolution. HeLa cells were harvested at early S phase, mid S phase and mitosis/G1 phase and either subjected to DMS treatment or nuclei were isolated and digested with DNase I. The purified DNA was restricted with DraI, separated in a sequencing gel, and electrotransferred to a nylon filter which was then hybridized with the DraI 5' upper probe (see Figure 3).

Several protein-DNA contacts were observed on the upper strand (sense strand) of the promoter, as indicated by the altered reactivity of guanine residues to DMS (Figure 4A). Strong enhancements were seen at position -195 and at one of the three guanine residues at positions -96 to -98. Partial enhancements were found at positions -240, -233, -221, -190, -85 and/or -84, and at position -74. Strong protection of guanine residues from DMS was observed at positions -215, -214, -213, -202, -199 and at one or two guanines out of the four at position -115 to -112. These data are consistent with the mapping of the light areas in the native genomic blotting autoradiographs, i.e., most of the guanine residues exhibiting enhancement or protection fall into the previously defined sites of protein-DNA interaction.

To localize the boundaries of the footprints on the upper strand, DNA samples from nuclei digested with DNase I were electrophoresed in parallel with the DMS samples (Figure 4B). Two regions of nuclease protection could be detected; site I was mapped to a region from -233 to -190 and site IIa from -145 to -122. Whereas site I was quite large (43 nucleotides), site IIa covered only about 23 nucleotides. Additionally, the protection around -200 (site I) was very strong suggesting tight DNA-protein interactions, while site IIa was less well protected--the DNase I pattern of the chromatin samples was only slightly different for site IIa from that of naked DNA. For that reason, we cannot be certain that the boundaries of the footprints are exact to the single nucleotide. We also detected an area of altered DNase I digestion that corresponded to site IIb in the native genomic blot (nt -96 to -74). While this area does not strictly meet the definition of a footprint, it is clear that the structure as detected by DNase I is different in chromatin than in deproteinized DNA and that several G residues in this area have enhanced reactivity to DMS in vivo (Figure 4A).

The DNase I pattern also displayed an interesting feature that had already been noted with the native genomic blotting: the very high accessibility of sequences between site I and site IIa to DNase I. The sensitivity was so high that it could be observed at single nucleotide resolution as a darker region in the chromatin samples when compared to naked DNA (see Figures 4B and 6).

The experiments that were performed for the upper strand of the H3

Figure 4: Fingerprints and footprints of the upper strand of the H3 histone gene promoter. The DNA was digested with DraI and hybridized with the DraI 5' upper probe (see Figure 3). **A.** DMS pattern of the upper strand. Lane 1: Guanine pattern of deproteinized DNA treated with DMS. Lanes 2-4: Guanine pattern of DNA from cells treated with DMS at early S phase, mid S phase and mitosis/G1 phase (1, 5 and 10 hrs after release from thymidine block, respectively). Open circles: guanine residues protected *in vivo*. Filled circles: guanine residues with enhanced reactivity *in vivo*. Half circles display only partial enhancement or protection. The sequence of the upper strand is displayed on the right and should be read from the bottom to the top. **B.** DNase I pattern. Samples were electrophoresed alongside those in A. Lanes 1-4: DNA isolated from nuclei of cells at mid S phase, and treated with 2, 3.5, 5 and 7.5 $\mu\text{g/ml}$ DNase I, respectively. Lanes 5-7: Deproteinized DNA digested with 0.1, 0.2 and 0.4 $\text{ng DNase I}/\mu\text{g}$ of DNA, respectively. The footprints are indicated on the right, with arrows marking their boundaries.

Figure 5: Fingerprints and footprints of the lower strand of the ST519 H3 histone promoter. The DNA was digested in A and D with HindIII and in B and C with NcoI. The probe for hybridization was the Nco 5' lower probe. A, B, and C. DMS pattern of the lower strand. Lane 1: Guanine pattern of deproteinized DNA treated with DMS. Lanes 2-4: Guanine pattern of DNA from cells treated with DMS at early S phase, mid S phase and mitosis/G1 phase (B and C only). Lane 5 (C): Guanine pattern of DNA from blocked cells treated with DMS. B and C display the same regions as A but from another gel with better resolution. The sequence of the lower strand is shown. Open and filled circles are the same as in Figure 4. D. DNase I pattern of the lower strand. Samples were electrophoresed alongside those in A. Lanes 1-3: Deproteinized DNA digested with 0.1, 0.2 and 0.4 ng DNase I/ μ g DNA, respectively. Lanes 4-7: DNA from nuclei of mid S phase cells treated with 2, 3.5, 5 and 7.5 μ g/ml DNase I, respectively. Arrows mark the limits of the footprints.

Figure 6: Effects of salt washes on the footprints of the upper strand of the H3 histone gene promoter. All samples are from early S phase cells. The DNA was digested with DraI and hybridized with the DraI 5' upper probe. The DNase I concentration in lanes 1-5 was 1-2 $\mu\text{g}/\text{ml}$. Lanes 1 and 2: DNA from nuclei treated with DNase I from two different cell cycle experiments. Lanes 3-5: DNA from nuclei that were washed with 0.18, 0.28 and 0.44 M NaCl prior to digestion with DNase I. Lanes 6-7: Deproteinized DNA digested with 0.15 and 0.3 ng DNase I/ μg DNA. Arrows mark the limits of the footprints.

promoter were repeated for the lower strand (Figure 5). In this case, the purified DNA from treated cells or nuclei was restricted with either HindIII (Figures 5A and 5D) or NcoI (Figures 5B and 5C) and then hybridized with the NcoI 5' lower probe (see Figure 3).

In the case of DMS treatment several sets of enhancement and protection of guanine residues could be identified. Five strong enhancements were seen at -238, -237, -235, -208 and -130. One partial enhancement was located at -146. Strong protection was observed at positions -200, -197 and -148, whereas one partial protection was detected at -135. Because some of the enhancements and protections were very subtle, it was necessary to analyze several different

gels and display different exposures in order to visualize these differences (Figures 5B and 5C). The protein contacts seen on the lower strand were localized in the same regions as on the upper strand.

To determine the limits of the footprints on the lower strand, DNA from nuclei digested with DNase I was electrophoresed in parallel with the DMS samples (Figure 5D). Due to technical limitations the area downstream of -120 could not be displayed on the lower strand: the HincII 3' upper probe (Figure 3) that we planned to use hybridized strongly to other H3 histone genes, indicating that the coding regions of different H3 genes are very conserved. However, sites I and IIa could be mapped at single nucleotide resolution, with site I covering a region from -229 to -184 and site IIa from -148 to -127. Therefore, site I covers a region of 45 nucleotides and site IIa 21 nucleotides. The protection of site I was staggered by 4 to 6 nucleotides with respect to the upper strand and site IIa was staggered by 3 to 5 nucleotides in the opposite direction. The footprint of site I was again obvious whereas the protection of site IIa was established only by a slightly different cutting pattern of DNase I in the chromatin lanes as compared to the naked DNA digestion, again suggesting weak interaction of the protein(s) with site IIa.

The patterns of DMS protection and enhancement on the two strands corresponded very well with the data from DNase I protection to give a picture of the staggered protein binding sites. The pattern of reactivity of guanine residues to DMS is different in this H3 histone gene than in the H4 gene we analyzed earlier (29). While we observed only protection of guanines from DMS in the H4 gene, this H3 gene displays a variety of different enhancements and protection with different strengths. In addition, the strength of protein-DNA interactions at each of the sites is reflected by the overall level of enhancement or protection.

Screening of sequences up to -430 bp did not reveal any other strong sites of protein-DNA interaction. Experiments are under way to determine if the DNase I sensitive band observed at -500 bp (Figure 1) is due to an additional protein-DNA interaction.

DNase I digestion pattern of salt extracted nuclei

To investigate the binding properties of these protein-DNA interactions, isolated nuclei from early S phase cells were washed with different salt concentrations prior to DNase I digestion. The results of these experiments provide an indication of the dissociation properties of the protein-DNA complexes (Figure 6). The DNase I footprints were partially abolished at a salt concentration of 0.18 M NaCl; some of the DNase I bands reappeared very

Figure 7: Summary of sites of protein-DNA interaction and contact points in the ST519 H3 histone gene in HeLa cells. The boxed area is the translated region of this gene. Symbols are the same as in Figures 4 and 5. Numbers above the sequence indicate positions with respect to the ATG codon. The horizontal arrow at -35 bp indicates the mRNA initiation site. Solid lines are derived from the genomic sequencing data (footprints). Dotted lines mark the protected areas found with the native genomic blotting method. The sequence is from Marashi, et al (48) with minor corrections that represent nucleotide differences between the pST519 sequence and HeLa genomic DNA.

faintly (Figure 6, lane 3). After the nuclei were washed with 0.28 M NaCl all of the DNase I bands were visible at the same intensity as the bands in the control lanes. Therefore, the proteins dissociate between 0.18 and 0.28 M NaCl, similar to the dissociation properties of the proteins bound to the H4 histone gene promoter.

DISCUSSION

To understand the functional properties of DNA sequences and regulatory molecules that influence the expression of human histone genes, we have used a

variety of techniques to investigate the chromatin structure and protein-DNA binding sites of a cell cycle dependent human H3 histone gene. The general accessibility of the coding and the flanking regions of this H3 histone gene to several nucleases reflects an actively transcribed gene. Additionally, one strong DNase I hypersensitive site is found in the proximal promoter region and a weaker sensitive site is located at -500 bp. These structural features strongly suggest that binding sites for regulatory molecules are located in the immediate 5' flanking region of this gene. Further analysis of the promoter region at high resolution revealed two, and possibly three, sites of protein-DNA interaction (site I: -230 to -185, site IIa: -145 to -122 and a tentative site IIb: -96 to -74) that may influence the structure and/or the transcription of this gene.

The comparison of nucleotide sequences and the contacts that the proteins have with specific guanine residues within sites I and II (Figure 7) lead us to predict that different proteins interact at each of these sites. Examination of the individual sites revealed the presence of three dyad symmetries associated with the specific protein contacts of site I, suggesting three proteins binding to this site. The first symmetry is in the 3' portion of site I and shows protection on both strands at the guanine residues forming

the symmetry. Immediately 3' to this region there is a strong enhancement of a guanine residue on the upper strand while 6 nucleotides upstream there is a strong enhancement on the lower strand.

A second symmetry can be found in the middle of site I where two

enhancements surrounding an asymmetric protection can be observed. The symmetry is not perfect since a C on the upper strand is replaced by a T on the lower strand.

Finally, a third symmetry is located adjacent to the 5' end of the DNase I protection of site I. It contains several enhancements, but no protected guanine residues:

Site IIa and the tentative site IIb, while each containing a CCAAT box motif, do not display the same contacts. Site IIa is followed at the 3' end by a GGGGOGGGGA sequence, which resembles the Sp1 core consensus sequence (34, GGGGOGGGGC) with two mismatches; however, we do not see the characteristic protection reflecting a bound Sp1 protein. Interestingly, in the H4 histone promoter we have analyzed previously (29) and which represents a more highly expressed histone gene, a sequence similar to the Sp1 core consensus sequence is intact and shows the characteristic protection for Sp1 that has been established *in vitro* (35).

Both site I and site IIa show protection or enhancement outside of the DNase I footprint. This may reflect the method of chromatin preparation: whereas the DMS treatment is done in whole cells, the DNase I treatment is performed on isolated nuclei. The isolation of nuclei might be sufficient to disrupt protein-DNA complexes, including possible attachment sites of the DNA with the nuclear matrix.

Comparison of the protein binding sites found for this H3 histone gene with the sites found recently in an H4 histone gene reveals that both sites I display very strong binding of the protein(s) to the DNA and have very high purine content on the upper strand (H4= 80% purines; H3= 70% purines). The proteins of both genes can be dissociated with sodium chloride exceeding 0.2 M and the sites do not show any differences throughout the cell cycle. Several dyad symmetries can be observed in site I of both genes, associated with specific DMS protection (H3 and H4) and specific DMS enhancements (H3). Although there is no extensive similarity between the two site I sequences, the octanucleotide AAATGACG is found in both the 3' part of H3 site I and the distal part of H4 site I; in both cases the G residues on the upper strand of this sequence display strong protection from reaction with DMS, raising the possibility of a shared protein-DNA interaction for these two genes. However, because of other differences in sequence and reactivities of G residues to DMS, it is unlikely that the DNA-binding proteins of these two genes are generally related. A comparison of the sequences of site I and IIa with the DNA binding domains of NF-1 (36), factor B (37), USF (38), MTLF (39), IgNF-A (40), IgNF-B (41), a factor binding to a adenovirus E1A inducible promoter (42), and the human beta interferon enhancer (43) did not reveal any significant similarities.

The strong interaction of proteins with site I leads us to the conclusion that this site is important for the expression of both the H3 and H4 histone genes. Because we have found site I, but not site II, protected in a heterologous system, when the H4 gene is expressed (unpublished data), this protein-DNA interaction might poise the gene for a basal level of transcription, and in conjunction with site II might determine the level of basal expression. Site II, which contains the histone specific sequence element GGTC (44,45) in both genes, might play a role in the modulation and rate of transcription of each individual histone gene. From earlier experiments it is known that this H3 histone gene is less transcribed than the H4 histone gene (unpublished observation), and we see weaker binding of the factors to site II in the H3 than the H4 gene. However, because the contacts at site II are so weak, we cannot exclude the possibility that other dynamic factors (e.g., RNA polymerase) (46) bind to these regions and thus we are simply detecting partial occupancy in vivo with our method. Additionally, although many H3 and H4 genes are found to be clustered (47), this is not the case for these particular H3 and H4 histone genes, which are found on two different chromosomes (unpublished data). Different expression and contacts may also be a reflection of the localization of these genes on different chromosomes.

The question then arises as to how these genes can be coordinately regulated to be transcribed at a higher level in early S phase. A few possible explanations are: A) There are dynamic protein-DNA (or protein-protein) interactions that are so subtle that they cannot be detected by the method employed; B) the proteins are modified at specific times of the cell cycle (phosphorylation, acetylation, poly ADP-ribosylation, etc.); C) additional sequences exist which confer cell cycle dependent expression to the gene and are located outside the areas we have examined; or D) the small region of shared sequence in site I represents the binding site for a protein involved in coordinating cell cycle transcription of these two genes. It will, therefore, be important to investigate the H2A and H2B genes at high resolution in order to generate a general model of how the coordinate expression of histone genes is regulated.

In summary, we have analyzed two human histone genes and found multiple in vivo protein-DNA interactions. These contacts suggest a complex pattern of factors that may regulate the expression of cell cycle dependent human histone genes.

ACKNOWLEDGEMENTS

We thank A. van Wijnen for many useful discussions. This study was supported by grants from the National Institutes of Health (GM 32010), National Science Foundation (PCM 83-18177) and March of Dimes Birth Defects Foundation (1-813). U.P. was supported by a fellowship from the Swiss National Foundation and S.C. by a National Research Service Award (GM 10453) from the NIGMS.

*To whom correspondence should be addressed at: Department of Cell Biology, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, MA 01655, USA

+Present address: University of Bern, Institute of Veterinary Virology, Laenggass-Strasse 122, CH-3012 Bern, Switzerland

REFERENCES

1. Plumb, M., Stein, J. and Stein, G. (1983) *Nucl. Acids Res.* **11**, 7927-7945.
2. Heintz, N., Sive, H. and Roeder, R.G. (1983) *Mol. Cell. Biol.* **3**, 539-550.
3. Baumbach, L., Plumb, M., Marashi, F., Stein, G. and Stein, J. (1984) *Biochemistry* **23**, 1618-1625.
4. Baumbach, L., Stein, G. and Stein, J. (1987) *Biochemistry* **26**, 6178-6187.
5. Schuemperli, D. (1986) *Cell* **45**, 471-472.
6. Stein, G.S., Sierra, F., Plumb, M., Marashi, F., Baumbach, L., Stein, J., Carozzi, N. and Prokopp, K. (1984) In Stein, G.S., Stein, J.L. and Marzluff, W.F. (eds.), *Histone Genes*. John Wiley and Sons, New York, 397-455.
7. Wu, C. (1980) *Nature* **286**, 854-860.
8. Stalder, J., Larsen, A., Engel, J., Dolan, M., Groudine, M. and Weintraub, H. (1980) *Cell* **20**, 451-460.
9. Elgin, S.C.R. (1988) *J. Biol. Chem.* **263**, 19259-19262.
10. Parish, R.W., Banz, E. and Ness, P. (1986) *Nucl. Acids Res.* **14**, 2089-2107.
11. Larsen, A., and Weintraub, H. (1982) *Cell* **29**, 609-622.
12. McGhee, J.D., Wood, W.I., Dolan, M., Engel, J.D. and Felsenfeld, G. (1981) *Cell* **27**, 45-55.
13. Wu, C. (1984) *Nature* **309**, 229-234.
14. Reeves, R. (1984) *Biochim. Biophys. Acta* **782**, 343-393.
15. Yang, J.Q., Remmers, E.F. and Marcu, K.B. (1986) *EMBO J.* **5**, 3553-3562.
16. Remmers, E.F., Yang, J.Q. and Marcu, K.B. (1986) *EMBO J.* **5**, 899-904.
17. Dynan, W.S. and Tjian, R. (1985) *Nature* **316**, 774-778.
18. Kroeger, P., Stewart, C., Schaap, T., van Wijnen, A., Hirshman, J., Helms, S., Stein, G. and Stein, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3982-3986.
19. Helms, S.R., van Wijnen, A., Kroeger, P., Shields, A., Stewart, C., Stein, J.L. and Stein, G.S. (1987) *J. Cell. Physiol.* **132**, 552-558.
20. van Wijnen, A.J., Stein, J.L. and Stein, G.S. (1987) *Nucl. Acids Res.* **15**, 1679-1698.
21. Dailey, L., Hanly, S.M., Roeder, G. and Heintz, N. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7241-7246.
22. Church, G. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991-1995.
23. Zinn, K. and Maniatis, T. (1986) *Cell* **45**, 611-618.
24. Southern, E.M. (1975) *J. Mol. Biol.* **98**, 503-517.
25. Pauli, U., Chrysogelos, S., Stein, J. and Stein, G. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 16-20.
26. Pauli, U., Chrysogelos, S., Stein, J. and Stein, G. (1988) *Biotechniques* **6**(2), 142-146.

27. Stein, G. and Borun, T.W. (1972) *J. Cell. Biol.* 52, 292-307.
28. Chrysogelos, S., Riley, D.E., Stein, G. and Stein, J. (1985) *Proc. Natl. Acad. Sci. USA* 82, 7535-7540.
29. Pauli, U., Chrysogelos, S., Stein, G., Stein, J. and Nick, H. (1987) *Science* 236, 1308-1311.
30. Chrysogelos, S., Pauli, U., Stein, G. and Stein, J. (1989) *J. Biol. Chem.* 264, 1232-1237.
31. Reed, K.C. and Mann, D.A. (1985) *Nucl. Acids Res.* 13, 7207-7221.
32. Roberts, G. and Wilson, G. (1985) *Focus (Bethesda Research Labs)* 7(3), 16.
33. Ephrussi, A., Church, G.M., Tonegawa, W. and Gilbert, W. (1985) *Science* 227, 134-138.
34. Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1986) *Science* 234, 47-52.
35. Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R. (1986) *Science* 232, 755-759.
36. Nagata, K., Guggenheimer, R.A. and Hurwitz, J. (1983) *Proc. Natl. Acad. Sci. USA* 80, 6177-6181.
37. Parker, C.S. and Topol, J. (1984) *Cell* 36, 357-369.
38. Sawadogo, M. and Roeder, R.G. (1985) *Cell*, 43, 165-175.
39. Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) *Cell* 43, 439-448.
40. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) *Nature* 319, 154-158.
41. Lenardo, M., Pierce, J. and Baltimore, D. (1987) *Science* 236, 1573-1577.
42. SivaRaman, L., Subramanian, S. and Thimmappaya, B. (1986) *Proc. Natl. Acad. Sci. USA* 83, 5914-5918.
43. Goodbourn, S., Burstein, H. and Maniatis, T. (1986) *Cell* 45, 601-610.
44. Hentschel, C.C. and Birnstiel, M.L. (1981) *Cell* 25, 301-313.
45. Wells, D.E. (1986) *Nucl. Acids Res.* 14, r119-r149.
46. Nick, H. and Gilbert, W. (1985) *Nature* 313, 795-798.
47. Sierra, F., Lichtler, A., Marashi, F., Rickles, R., VanDyke, T., Clark, S., Wells, J., Stein, G. and Stein, J. (1982) *Proc. Natl. Acad. Sci. USA* 79, 1795-1799.
48. Marashi, F., Helms, S., Shiels, A., Silverstein, S., Greenspan, D.S., Stein, G. and Stein, J. (1986) *Biochem. Cell Biol.* 64, 277-289.