Volume 15 Number 4 1987 Nucleic Acids Research

A nuclear protein with affinity for the 5' flanking region of a cell cycle dependent human H4 histone gene *in vitro*

Andre J.van Wijnen, Janet L.Stein and Gary S.Stein

Department of Biochemistry and Molecular Biology, Department of Immunology and Medical Microbiology, University of Florida College of Medicine, Gainesville, FL 32610, USA

Received September 19, 1986; Revised and Accepted January 12, 1987

ABSTRACT

A nuclear protein with affinity for the 5' flanking region of a cell cycle dependent human H4 histone gene has been partially purified from nuclear extracts of human HeLa S3 cells. The region involved in the binding of the protein has been localized to an upstream DNA segment using an electrophoretic mobility shift assay. This DNA segment is devoid of RNA polymerase II consensus sequences and contains both homopurinic and A/T rich tracts. Analogous experiments have identified a similar, and perhaps identical, factor that has affinity for a cell cycle dependent human H3 histone gene promoter. This protein appears to bind to a DNA segment containing A/T rich sequences that bear homology with the binding region of the H4 histone promoter. Cell synchronization experiments have shown that the overall affinity of the protein(s) for the H3 and H4 histone 5' flanking regions in vitro is not dramatically altered during the cell cycle. Although the rate of histone gene transcription is modulated during early S phase, transcription occurs throughout the cell cycle. Hence, the protein(s) we have detected here may play a role in the basal expression of these genes.

INTRODUCTION

Human histone genes represent a heterogenous, clustered, multi-gene family (1-4) with a complex chromosomal organization (5,6). Cell cycle dependent human histone genes are coordinately expressed in tight conjunction with DNA synthesis and their expression is regulated at both transcriptional and post-transcriptional levels (reviewed in 7-9). The transcriptional component involves a 3 to 5 fold, transient increase in the rate of histone mRNA synthesis at the onset of S phase, which rapidly declines to a basal rate that persists throughout the cell cycle (10). This process has been found to be paralleled by a sequential remodeling of chromatin architecture in the 5' regulatory region of a cell cycle dependent human H4 histone gene, designated FO108 (11). This FO108 H4 histone gene exhibits reversible cell cycle dependent alterations in both nucleosomal organization (12) and sensitivity towards the nucleases S1 and DNase I (13).

DNA sequence requirements for histone gene transcription have been studied

in some detail in our laboratory by using the FO108 H4 histone gene and a cell cycle dependent H3 histone gene, designated ST519 (14) as model systems. Transcription experiments with the FO108 H4 histone gene in vitro, employing whole cell extracts from exponentially growing HeLa cells, have shown that only a limited set of sequences (encompassing a TATA-box at position -60 from the coding region) is required for accurate transcription initiation of this gene (11). These results were later confirmed by Heintz and collaborators with S phase specific nuclear extracts, using an independently isolated H4 histone gene designated Hu4A (15,16). These workers also showed that sequences further upstream (between nucleotides -70 and -110 of this particular gene) might be involved in modulating the efficiency of transcription in vitro (16).

The H4 histone gene FO108 and the H3 gene ST519 have been tested in vivo in transient expression systems (14,17). The FO108 H4 histone gene has also been tested in vivo in stable, integrated (20) and episomal systems (18). These experiments have revealed that 5' flanking regions of both human histone genes can function as promoters in vivo and apparently contain enough DNA sequence information for the specific recognition by putative trans-acting factors of both primate (14) and murine (14,17,18) origin; a similar conclusion has been drawn for the Hu4A histone H4 gene when chromosomally integrated in murine cell lines (19). In addition, it has been shown for the FO108 H4 gene, that a minimal promoter region spanning nucleotides up to position -240 (with respect to the H4 coding sequence) is sufficient for accurate transcription initiation in vivo (17,18). Results obtained by assaying expression levels of 5' deletion mutants of the FO108 H4 gene in stable murine cell lines furthermore have given an indication for positive cis-regulatory control elements in the 5' flanking region of this gene upstream of position -240 (20).

Sequence specific DNA binding proteins acting on cis-regulatory control elements have been hypothesized to be key elements in eukaryotic gene transcription (21), and considerable progress has been made in the isolation of such proteins (22-26). Yet, DNA binding proteins with affinity for the 5' regulatory regions of cell cycle dependent histone genes have not been identified. Characterization of such proteins is crucial in understanding the mechanisms underlying the transcriptional control of histone gene expression. We have employed the electrophoretic mobility shift assay (gel retardation assay) (27,28) in order to isolate DNA binding proteins with affinity for a human H4 histone gene regulatory region in vitro. In this paper we report the

partial purification of a factor that has apparent affinity for an upstream DNA segment in the 5' flanking region of the FO108 H4 histone gene. In addition, we have found a similar and perhaps identical factor that associates with a DNA fragment within the 5' regulatory region of an H3 histone gene.

MATERIALS AND METHODS

DNA Probes

The cell cycle dependent core human H3 and H4 histone genes used in this study were derived from two genomic histone gene clusters cloned into lambda Charon 4A (1). The upstream flanking regions of both an H4 gene (FO108) (11) and an H3 gene (ST519) (14) were subcloned in plasmid vectors and used to generate three probes. Two of these probes (H4-A and H4-B) were derived from the 5' flanking region of the H4 gene FO108, and the third probe (H3-A) comprises the immediate upstream flanking sequences of the H3 gene ST519 (see Figure 1). Probe H4-A was generated by a SmaI/BanII restriction endonuclease digestion of the plasmid subclone pFO002 (20) giving rise to a 223 bp fragment starting 13 nucleotides upstream of the H4 coding sequence. This probe covers the functional H4 gene promoter as defined in vitro and in vivo (11,17) and a region hypersensitive to S1 nuclease and DNase I (13) (see Figure 1). Probe H3-A was prepared by double digestion of plasmid pST519ΔH (14) with HpaI/HindIII which liberated a 190 bp fragment spanning the analogous region for the H3 gene. Probe H4-B was derived by digesting subclone pFO005 (20) with Eco RI and comprises nucleotides -244 to -445, which are located immediately upstream of probe H4-A; its size is 242 bp due to sequences derived from the pUC13 polylinker.

Probe fragments were dephosphorylated using calf intestine alkaline phosphatase, end-labelled with γ-^{32}P-ATP using T4 polynucleotide kinase and purified by polyacrylamide gel electrophoresis according to standard procedures (29).

Nuclear Extracts

Nuclear extracts were prepared from exponentially growing HeLa S3 cells essentially according to the procedure of Challberg and Kelly (30). Mechanically lysed cells (5 X 10^8) were separated into cytosol and crude nuclei by low speed centrifugation (1500 g) for five minutes. The cytosol supernatant was dialyzed against buffer A (25mM Hepes-NaOH, pH 7.5, 20mM NaCl, 1mM DDT, 1mM EDTA, 20% glycerol) and stored frozen at -70°C. The fresh nuclear pellet was extracted with a sucrose buffer (50 mM Hepes-NaOH, pH 7.5, 10% sucrose, 400mM NaCl) followed by centrifugation at 12,000 g for ten

minutes. The supernatant was dialyzed against buffer A and the resulting solution, representing the crude nuclear extract, was stored at -70°C. Buffer solutions in all the above and subsequent steps contained 0.5 mM phenyl methyl sulfonyl fluoride (PMSF) and 10 mM sodium bisulfite. Protein estimation was performed according to the method of Bradford (31) with bovine serum albumin as the standard.

Fractionation of Nuclear Extracts

Crude nuclear extracts from 4 X 10^9 cells were centrifuged at 100,000 g (31) for forty minutes in order to remove remaining cytoplasmic debris and the salt concentration was adjusted to 0.2 M KCl. The resulting solution (25 ml) was applied to a DEAE-Sephacel column (Pharmacia; dimensions 2.3 X 11 cm) equilibrated with buffer A containing 0.2 M KCl, and the flow-through was collected (D200 fraction; 30 ml). Proteins more tightly associated with the DEAE column were eluted with 15 ml of buffer A containing 1.0 M KCl (D1000 fraction).

The D200 fraction was further separated into three components by passage over a 15 ml phosphocellulose column (P11, Whatman) equilibrated with buffer A containing 0.04 M KCl and 0.01% Nonidet P40 and by sequential elutions at 0.04 M KCl (20 ml), 0.35 M KCl (20 ml) and 1.0 M KCl (15 ml) in the same buffer. Protein-containing fractions, designated P40, P350 and P1000, respectively, for each salt step, were pooled and dialyzed against buffer A containing 0.01% Nonidet P40.

Electrophoretic Mobility Shift Assay

The electrophoretic mobility shift assay (gel retardation assay) (27, 28) was basically performed according to Carthew et al. (23). Binding reactions typically contained 2 ng of probe, 1 µg of competitor DNA (sonicated E. coli DNA or poly(dI-dC).poly(dI-dC)) and the appropriate amounts of protein (1-5 µg) in a binding buffer (12 mM Hepes-NaOH, pH 7.9, 12% glycerol, 60 mM KCl, 5mM $MgCl_2$, 4mM Tris-HCl, 0.6 mM EDTA, 0.6 mM DTT) in a total volume of 20 µl. Samples were incubated for thirty minutes at 30°C and electrophoresed in a 4% low ionic strength polyacrylamide gel (acrylamide:bisacrylamide 30:1). Gels were dried and subjected to autoradiography.

Deletion analysis using the gel retardation assay was performed by sequential and reciprocal shortening of a relatively large probe (200 bp or more) by various restriction enzymes ("stairway assay"). In brief, 20 ng aliquots of a single end-labelled probe were incubated with excess amounts of the restriction enzyme(s) of choice and the reactions stopped by extraction with phenol/chloroform. The DNA was then precipitated with ethanol in the

presence of 10 µg of competitor DNA to ensure optimal recovery of both probe and competitor DNA. The shortened fragments were mixed with the same amount of protein and simultaneously electrophoresed on a polyacrylamide gel, showing a stairway-like pattern upon autoradiography.

Cell Synchronization

Synchronization of HeLa S3 cells by double thymidine block was performed according to well established procedures (10,12,13,18,32). DNA synthesis rates after release from thymidine block were monitored by ^{3}H-thymidine pulse labelling experiments. ^{3}H-Thymidine incorporation into TCA insoluble material increased 10-fold during the initial 4-5 hours after release and declined to its lowest point after 8 hours (data not shown).

DNase I footprinting

DNase I protection analysis was carried out by incubating a 5-fold scaled up binding reaction (100 µl; containing E. coli DNA) with 1 u of DNase I (Sigma) for 60 seconds at 20°C. The reaction was stopped by adding 20 mM EDTA, 0.1% SDS and 10 µg E. coli DNA to the sample on ice. Subsequently,

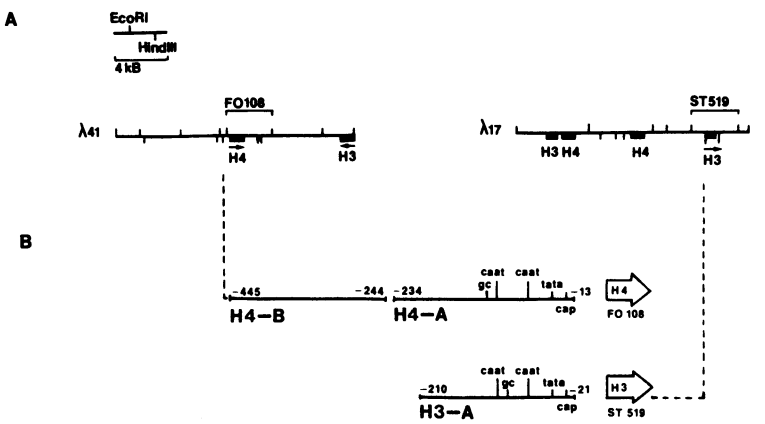


Figure 1: (A) Restriction maps of the human histone gene clusters λHHG41 and λHHG17 (1) used for the generation of probes (see text for details). Arrows indicate the direction of histone mRNA transcription. (B) Schematic drawing of the 5' flanking DNA fragments from the FO108 H4 histone gene (H4-A and H4-B) and the ST519 H3 histone gene (H3-A) used as probes. Dashed lines indicate the location in the gene clusters from which the restriction fragments were derived. Numbers indicate the positions of probe termini with respect to the histone coding sequences (symbolized by open arrows). The locations of DNA sequences homologous to the TATA-, CAAT- and GC-boxes (21) and the histone mRNA cap site (11,17) are indicated by the corresponding abbreviations. A DNase I/S1 hypersensitive region (see Figure 6) is present in the H4-A probe between position -234 and the GC-box (13).

samples were extracted with phenol/chloroform, ethanol precipitated, and analyzed on a denaturing 6% polyacrylamide, 6 M urea gel.

RESULTS AND DISCUSSION

Proteins from Unfractionated HeLa Cell Nuclear Extracts Bind to Human H4 Histone Gene 5' Flanking Sequences

The gel retardation assay has been used in this study to detect factors, derived from a HeLa nuclear extract, that bind to the 5' flanking region of a human H4 histone gene (FO108). This assay, which detects the retarded migration rates of DNA fragments bound to proteins, is based on increased stability of these complexes in low ionic strength polyacrylamide gels (33) and has become an effective tool for the isolation of eukaryotic DNA binding proteins (for example 23-25). Three probes have been employed; one spans the H4 promoter region (H4-A, containing nucleotides [nt] -13 to -234), the second comprises the DNA sequences directly upstream from the first probe (H4-B, nt -244 to -445), and the third (H3-A) encompasses the promoter region of a human H3 histone gene (ST519). Probe H3-A (nt -21 to -210) is analogous in size and position to probe H4-A (see Figure 1).

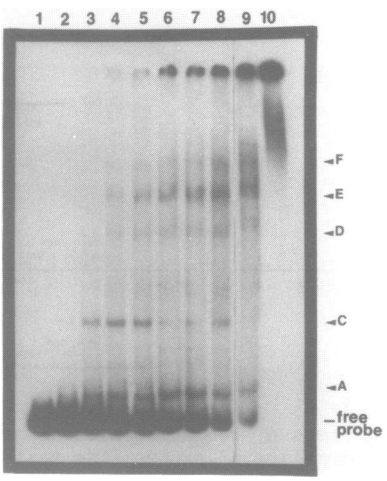


Figure 2: Detection of proteins with affinity for the H4 promoter region (nt -13 to -234) in an unfractionated nuclear extract using the gel retardation assay. Prominent DNA-protein complexes (retarded probe bands) are indicated by capital letters (A,C,D,E and F). Ln 1: H4-A probe (2 ng) and E. coli DNA (1 µg) only. Ln 2-10: addition of increasing amounts of nuclear extract protein; respectively, 0.1, 0.4, 0.8, 1.2, 1.6, 2.0, 2.4, 4.0 and 8.0 µg of protein. Not indicated is a subband originally designated B, which was not consistently observed in all experiments using the H4-A probe.

Nuclear extracts from exponentially growing HeLa S3 cells were prepared essentially according to the protocol originally published by Challberg and Kelly (30), which has been modified and extended by several others (34 and references therein). This extraction protocol yielded the first sequence specific DNA binding protein (NFI) isolated from human cells (35). In an initial experiment, various amounts of nuclear extract were incubated with the radiolabelled H4-A promoter fragment in the presence of excess amounts of sonicated E. coli DNA to suppress non-specific DNA binding, and the DNA/protein mixtures were subjected to polyacrylamide gel electrophoresis. At least five probe bands with altered migration rates were observed (Figure 2) and indicate the presence of bound factors causing this mobility shift.

Similar experiments in which both the H4-A promoter probe (nt -13 to -234) and the H4-B upstream probe (nt -244 to -445) were assayed in parallel reactions revealed that the probes each give rise to a different bandshifting

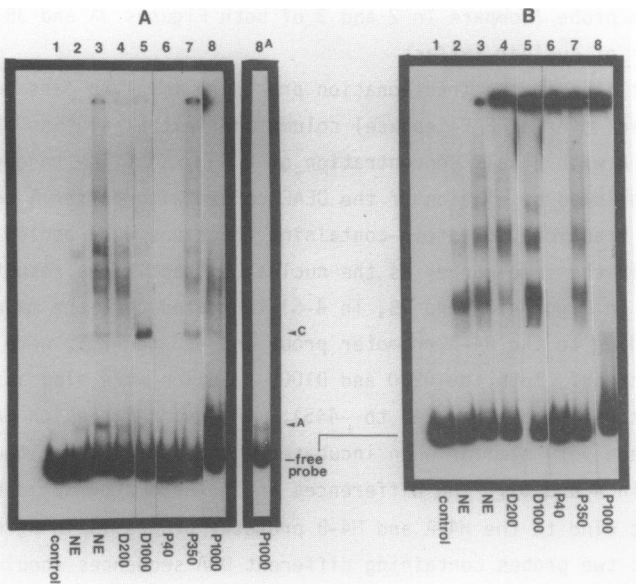


Figure 3: Monitoring of column fractions for proteins binding to the H4-A probe (nt -13 to -234) (panel A) and the H4-B probe (nt -244 to -445) (panel B) using the gel retardation assay. Ln 1: probe DNA (2 ng) and E. coli DNA (1 µg) only. Ln 2-8: Proteins added to the binding reaction. Ln 2-3: 0.8 µg and 2 µg nuclear extract protein, respectively. Ln 4: 2 µg D200 protein; ln 5: 0.5 µg D1000 protein. Ln 6: 6 µg P40 protein ; ln 7: 2 µg P350 protein; ln 8: 1 µg P1000 protein. Ln 8A: 0.5 µg P1000 protein (shorter exposure of an analogous experiment). Indicated by capital letters are an H4-A specific subband (designated A) and an apparently non-specific subband (designated C).

pattern with respect to relative migration rates and intensities of the subbands (for instance, Figure 3A and 3B; lanes [ln] 2 and 3). In order to confirm that the factors mediating the probe retardation were proteins, the nuclear extract was preincubated with proteinase K or 0.1% sodium dodecyl-sulfate (SDS) or heated for ten minutes at 100°C; in none of these cases were probe fragments retarded in the assay (data not shown). In addition, we found that most of the proteins that bound to the H4-A probe were enriched in the nuclear extract when compared with a cytosol preparation (data not shown).

The latter two findings, together with the observation that two probes of approximately equal length displayed different gel retardation patterns, suggested the presence of nuclear proteins with specific affinity for the different DNA sequences within the probe. We therefore initiated fractionation of these proteins using liquid chromatography. Initially we focused on a protein that gave rise to a fast migrating DNA/protein complex (designated A in Figures 2 and 3A), with affinity for the H4-A promoter probe but not the H4-B upstream probe (compare ln 2 and 3 of both Figures 3A and 3B).

Fractionation of Nuclear Extract

The first step in the fractionation procedure involved passage of the nuclear extract over a DEAE-Sephacel column and initial elution of proteins using buffer A with a salt concentration of 0.2 M KCl (flow through; D200 fraction), followed by elution of the DEAE column with buffer A containing 1.0 M KCl (D1000 fraction). Protein-containing fractions were pooled, dialyzed and assayed in the same manner as the nuclear extract. The results of this assay (shown in Figures 3A and 3B, ln 4-5) indicated that the majority of proteins binding to the H4-A promoter probe (nt -13 to -235) were present in the D200 fraction. Both the D200 and D1000 fraction were also assayed using the H4-B upstream probe (nt -244 to -445). The gel retardation patterns of this probe were very similar when incubated with each of these two fractions (Figure 3B, ln 4 and 5). The differences in the chromatographic behavior of proteins that bind to the H4-A and H4-B promoter probes is in agreement with the idea that two probes containing different DNA sequences should interact with different factors.

A predominant probe DNA/protein complex (designated C, Figure 3A, ln 5) was observed when the D1000 fraction was used in the gel retardation assay. Preliminary mapping experiments using restriction site deletion analysis in combination with the gel retardation assay (described in detail below), indicated that this protein lacked affinity for any particular H4-A probe segment (data not shown). That is, the protein was found to be associated

with the H4-A probe regardless of its size when shortened by restriction enzymes from either end; also, no loss in intensity of the shifted band was observed. Another property of this factor was that it was relatively abundant in a cytosol preparation (data not shown) and that the migration rate of the complex was relatively independent of the size of the probe. These data suggest that this factor which caused a large shift in the migration of the probe band is a large cellular protein with non-specific DNA binding activity.

The second fractionation procedure included a three step salt elution from a phosphocellulose column of the D200 fraction, which contained the majority of proteins that associated with the H4-A promoter probe. Fractions eluting at salt concentrations of 0.04 M, 0.35 M and 1.0 M KCl (designated P40, P350 and P1000, respectively) were separately pooled and assayed with both the H4-A and H4-B probes (ln 6-8, Figure 3A and 3B). We observed that the fastest migrating DNA/protein complex (band A) could not be formed when the H4-A probe was incubated with either the P40 (6 µg protein added) or the P350 fractions (2 µg proteins added); however, an intense A band was detected when the P1000 fraction (1 µg protein) was used in the binding reaction (Figure 3A, ln 8 and 8A). Lane 8 in Figure 3A shows the slower migrating bands of low intensity more clearly; the A-band however, is overexposed and is no longer distinguishable from the free probe. A lighter exposure of an analogous experiment, with only 0.5 µg of the P1000 fraction added, demonstrated more convincingly the presence of the A-band (Figure 3A, ln 8A). These results indicated that the protein causing the slight retardation of the H4-A promoter probe, but not the H4-B upstream probe, eluted from the phosphocellulose column with salt concentrations between 0.35 M and 1.0 M KCl.

Competition binding studies

The results described thus far were obtained by using E. coli DNA as competitor-substrate for non-specific DNA binding proteins; but several titration experiments, in which a constant amount of protein was incubated with increasing amounts of E. coli DNA, showed general competition of all the H4-A specific subbands by this type of DNA (see, for example, Figure 4A; ln 2-7). To suppress this non-specific competition we used the double-stranded, alternating co-polymer poly dI/dC (23). In general, increasing amounts of this competitor DNA showed less decrease in the intensity of specific subbands (Figure 4B, ln 5-7; e.g., bands D,E,F), as previously reported (24). However, the retarded probe band designated A was largely quenched by the addition of low amounts (100 ng) of poly dI/dC.

Although the basis for this differential competition phenomenon has not

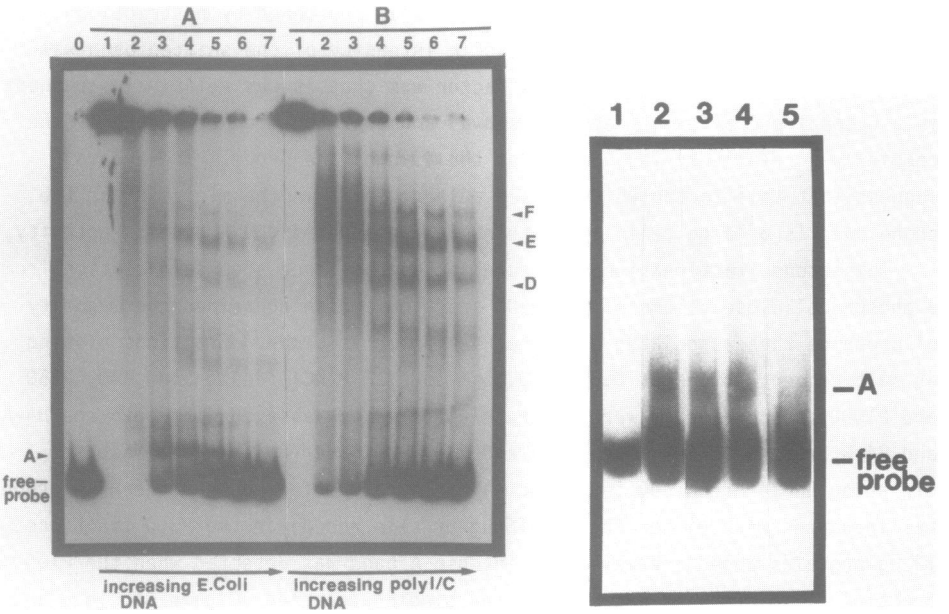


Figure 4: Titration experiment using E. coli DNA (part A) or poly dI/dC (part B) as competitors in the gel retardation assay. Ln 0: H4-A probe DNA (2 ng) only. Ln 1-7 (part A and B): H4-A probe DNA (2 ng) and 2 µg of nuclear extract protein with increasing amounts of E. coli DNA (A) or poly dI/dC (B). Ln 1-7: respectively, 0, 0.1, 0.2, 0.4, 1.0, 2.0 and 4.0 µg of competitor DNA. Prominent DNA-protein complexes are indicated by capital letters (A,D,E, and F). Part C: Competition experiment using pUC8 DNA (ln 2,3) and pF0002 DNA (ln 4,5). Ln 0: H4-A probe DNA (2 ng) and E. coli DNA (1 µg) only; ln 2-5: 1 µg P1000 protein added. Ln 2 and 4: 100 ng pDNA; ln 3 and 5: 2 µg pDNA added. Salt concentration was raised to 100 mM KCl in the binding reaction.

yet been tested, the specific competition of a DNA binding protein by synthetic polymers such as poly dI/dC is not without precedent. Mouse α protein (36), which recognizes A/T rich DNA sequences, is effectively impaired in binding to its radiolabelled cognate sequences in the presence of poly(dA-dT).poly(dA-dT) or poly(dI-dC).poly(dI-dC) but to a lesser extent in the presence of sonicated E. coli DNA or poly(dG-dC).poly(dG-dC).

Competition experiments were also performed using supercoiled plasmid DNA that carries an insertion encompassing the H4-A probe fragment (pF0002); in these experiments, equivalent amounts of supercoiled vector DNA (pUC8) were used as a control (Figure 4C). Incubation of a standard mixture, containing 1 µg E. coli DNA and 2 ng H4-A probe, with 100 ng or 2 µg of vector DNA alone did not result in competition of the A-band (Figure 4C; ln 2,3). However,

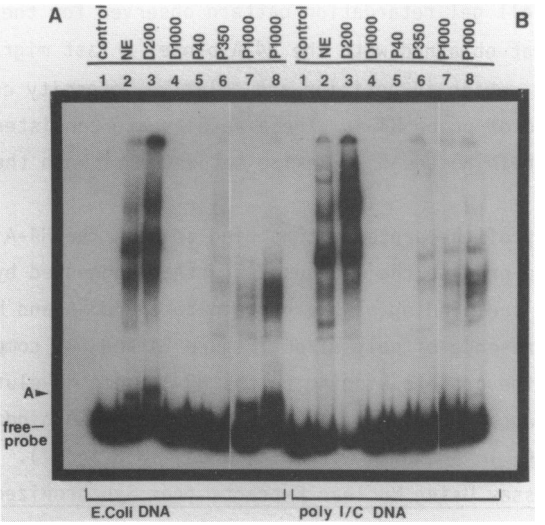


Figure 5: Monitoring of column fractions for proteins binding to the H3-A promoter probe (nt -21 to -210) using either E. coli DNA (part A) or poly dI/dC (part B) as competitor DNA. Ln 1: H3-A probe DNA (2 ng) and competitor DNA (1 µg) only. Ln 2-8: Nuclear extract protein added to the binding reaction. Ln 2: 2 µg of nuclear extract protein; ln 3: 4 µg of D200 protein; ln 4: 1 µg D1000 protein. Ln 5: 6 µg of P40 protein; ln 6: 0.9 µg of P350 protein; ln 7: 1 µg of P1000 protein and ln 8: 2 µg of P1000 protein. The slightly retarded probe subband, visible only when using E. coli DNA as competitor is indicated by the letter A.

incubation with 2 µg of pF0002 DNA, corresponding to a 40:1 molar ratio of the unlabeled H4-A probe fragment (inserted in pF0002) versus the radiolabeled H4-A probe, resulted in a profound decrease in intensity of the A-band (Figure 4C, ln 5). These results show that the A-band is competed by the insertion spanning the H4-A probe fragment and suggest that the A-band protein binds specifically to this fragment.

Detection of a similar protein binding to the H3 5' regulatory region

We then investigated the possibility that proteins similar to those which bind to the human H4 histone gene promoter would interact with an analogous region of a cell cycle dependent human H3 histone gene promoter. The probe used was the H3-A fragment (Figure 1) that spans positions -21 to -210 and is comparable in size to the H4-A probe (see Figure 5). The same protein fractions and quantities as used for experiments with the H4-A probe were assayed. In addition, the gel retardation assay was performed in the presence of E. coli DNA (Figure 5A; ln 1-8) or poly dI/dC (Figure 5B; ln 1-8).

Although the overall gel retardation pattern observed for the H3-A probe was different from that obtained with the H4-A probe, a fast migrating subband (A-band) was seen, which in position and relative intensity corresponded with the shifted A-band of probe H4-A. These results are consistent with binding of the A-band protein to the 5' flanking sequences of both the H4 and H3 human histone genes.

The similarity of the proteins that bind to both the H4-A and H3-A promoter probes to produce the A-bands is further suggested by the following observations: (I) the binding of the protein to the H3-A and H4-A probes is impaired in the presence of poly dI/dC (Figure 5A and 5B, compare ln 2, 3, 7 and 8); and (II) the protein binding to the H3-A probe co-elutes over two different chromatography columns with the H4-A probe DNA binding protein (compare Figure 3A, ln 4 and 8, with Figure 5A, ln 3 and 7).

Gel Retardation Assay Using Nuclear Extracts from Synchronized Cells

The possibility that the protein(s) binding to our probes would have a DNA binding activity modulated at different stages during the cell cycle was explored by preparing nuclear extracts of synchronized HeLa cells at various times after release from a double thymidine block. Both H3-A and H4-A probes were incubated with identical increasing amounts of nuclear extract protein derived from cells at 1hr, 5 hr, and 10 hr after release, representing respectively the onset of S phase, the peak of S phase and the mitosis + G1 period. These results revealed that the protein(s) responsible for the fast migrating A-band is extractable and detectable at all three time points examined, indicating that the overall DNA binding activity of the protein(s) is not tremendously altered during the cell cycle (data not shown). This could suggest that the protein(s) might not be directly involved in early S phase specific enhancement of histone mRNA transcription, but does not rule out involvement of the protein in rendering the genes transcribable at a basal level throughout the cell cycle. This result is consistent with *in vivo* data obtained in our laboratory (U. Pauli, S. Chrysogelos, G. Stein, J. Stein and H. Nick, manuscript submitted) that suggest the presence of several factors bound to the H4 histone promoter in intact cells throughout the cell cycle.

Deletion Analysis Using the "Stairway Assay"

The question as to whether the protein(s) detected had affinity for particular segments of the H3-A or H4-A probe was addressed by analyzing the binding of protein fractions to probes that were progressively shortened from either direction. The shortened probes were simultaneously analyzed in a single polyacrylamide gel in the presence of increasing amounts of protein,

giving a characteristic stairway pattern for both the probes and their subbands ("stairway-assay"; see Materials and Methods).

In this approach, the H4-A probe was ^{32}P-labelled at the 5' terminal SmaI site and small aliquots of the probe were digested with various restriction enzymes. Electrophoresis of these smaller sized probes in the presence of either crude nuclear extract or purified protein fractions (D200 or P1000)

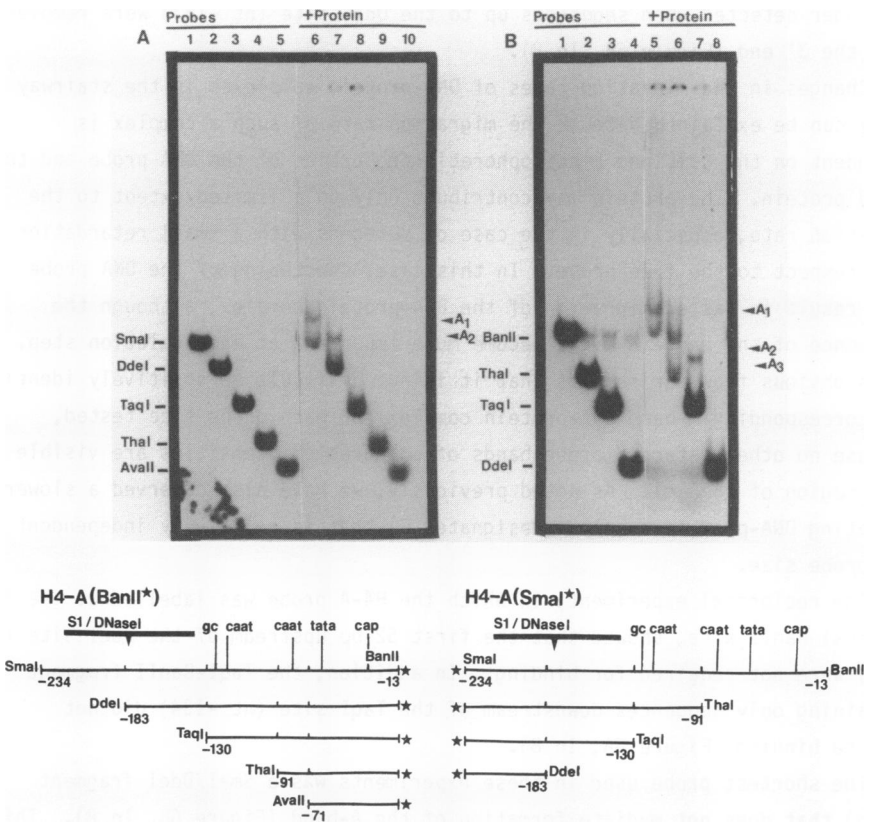


Figure 6: Restriction site deletion analysis of the H4-A probe using the gel retardation assay. The H4-A probe is labeled at the 5' terminal SmaI site (panel B) or 3' terminal BanII site (panel A) (see text for details). Lower panel: schematic drawings of the probe (see also Fig. 1B) when shortened by the indicated restriction enzymes. A black star indicates the position of the ^{32}P label. Panel A (ln 1-5) and panel B (ln 1-4): shortened probes and E. coli DNA only (1 µg total DNA). Panel A (ln 6-10) and panel B (ln 5-8): addition of 4 µg of D200 protein. The positions of the shifted probe A-band are indicated by arrows (A1, A2, A3). The faint bands in panel B (ln 2-4) are a result of incomplete restriction endonuclease digestion of the H4-A probe.

revealed changes in the migration rates of the protein-DNA complexes as a function of the probe size without visible loss in intensity of each subband (designated A1, A2 and A3; Figure 6). More importantly, the absence of specific gel bands upon removal of particular DNA sequences indicated that some protein-DNA complexes could no longer be formed (Figure 6A and 6B, ln 8). The complex could still be formed when sequences at the 3' end of the H4-A probe were deleted up to the TaqI site (nt -130), removing all RNA polymerase II consensus sequences (TATA-; CAAT- and GC-box) (21); however the A-band was no longer detected when sequences up to the DdeI site (nt -183) were removed from the 3' end (Figure 6B, ln 8).

Changes in the migration rates of DNA-protein complexes in the stairway assay can be explained because the migration rate of such a complex is dependent on the combined electrophoretic properties of the DNA probe and the bound protein. The protein may contribute only to a limited extent to the migration rate, especially in the case of subbands with a small retardation with respect to the free probe. In this case, shortening of the DNA probe will result in faster migration of the DNA-protein complex, although the influence of the protein might become more important at each deletion step. It is obvious from our results that it is not difficult to positively identify the corresponding A-band DNA-protein complex for each probe size tested, because no other retarded probe bands of comparable intensities are visible in this region of the gel. As noted previously, we have also observed a slower migrating DNA-protein complex (designated C) that is relatively independent of the probe size.

The reciprocal experiment, in which the H4-A probe was labelled at the 3' terminal BanII site, showed that the first 52 bp upstream of the DdeI site (nt -183) were not required for binding. In addition, the TaqI-BanII fragment containing only sequences downstream of the TaqI site (nt -134) did not mediate binding (Figure 6A, ln 8).

The shortest probe used in these experiments was a SmaI/DdeI fragment (48bp) that does not mediate formation of the A-band (Figure 6B, ln 8). This result is most likely not a reflection of only the probe size. Firstly, because in the reciprocal experiment it has been shown that this fragment is dispensible for formation of the A-band. Secondly, because binding of other proteins to this fragment was observed when employing crude nuclear extracts in the stairway assay (unpublished observations). Moreover, probe fragments consisting of only 35bp have been found to be effective in the detection of DNA binding proteins in the gel retardation assay (37). Because of these

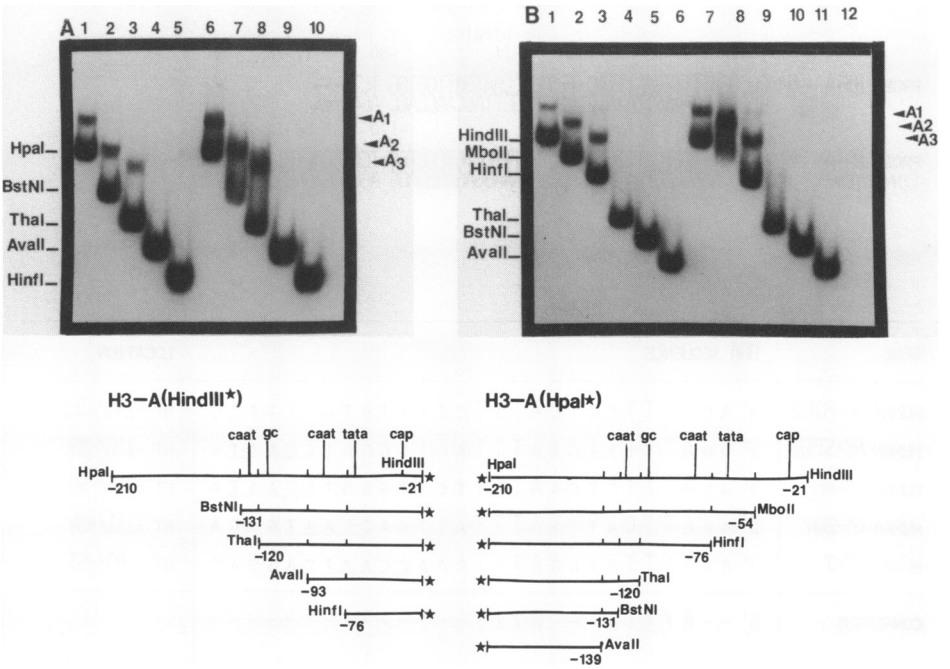
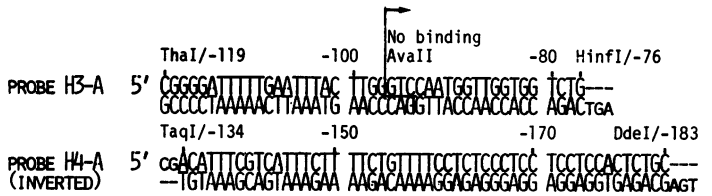


Figure 7: Restriction site deletion analysis of the H3-A probe using the gel retardation assay. The H3-A probe is labeled at the 5' terminal HpaI site (panel B) or 3' terminal HindIII site (panel A) (see text for details). Lower panel: Schematic drawings of the probe (see also Fig. 1B) when shortened by the indicated restriction enzymes. A black star indicates the position of the ^{32}P label. Panel A (ln 1-5) and B (ln 1-6): shortened probes and E. coli DNA (1 µg total DNA) in the presence of 0.5 µg of P1000 protein; panel A (ln 6-10) and panel B (ln 7-12): same, but 1 µg P1000 protein added. The positions of the shifted probe A-band are indicated by arrows (A1, A2,A3).

considerations and the observation that the data from this reciprocal assay are compatible, we have localized a putative protein binding domain within DNA sequences between the DdeI and TaqI sites, located at positions -183 and -134 with respect to the H4 histone coding sequence.

The analogous experiment was performed with the H3-A probe labelled at either the 5' terminal HpaI site or the 3' terminal Hind III site. The data show clearly that DNA sequences upstream of the TaqI site at -119 (Figure 7A, ln 3) and downstream of the HinfI site at -76 (Figure 7B, ln 3) are dispensable; yet a fragment containing only sequences downstream of an AvaII site at position -94 is no longer capable of binding (Figure 7A, ln 4). This establishes the putative protein binding domain in the H3-A probe between



GENE	DNA SEQUENCE	LOCATION
HUMAN H4-FO108	5' A C A T T T C G T C A T T T C T T T T C T G T T T T C C	NT -134/-161
HUMAN H3-ST519	5' G G A T T T T T G A A T T T A C T T G G G T C C A A T G	NT -116/-89
HUMAN H3-5B	5' G G A T T T T T G A A T T T T C T T G G G T C C A A T A	NT -116/-89
HUMAN H3-26H	5' A A A T T C A T T G A T T T A T T G A C C A A T G A G A	NT -111/-84
MOUSE H3-2	5' G C A T T T G A A G A T T T C A A C C A A T C A G G A G	NT - 90/-63
CONSENSUS	5' - - A T T T - - - - A T T T - - - - - - - - - - - - -	

Figure 8: (Upper panel) DNA sequences of the restriction fragments of probe H3-A and H4-A involved in binding of the A-band protein. The H4-A probe is shown in inverted orientation to facilitate DNA sequence comparison between the two probe segments. Lower case letters represent single-stranded protruding ends resulting from restriction enzyme digestion. An arrow pointing downstream in the H3-A sequence indicates the position of an AvaII site; a fragment containing only sequences from this site to the downstream HindIII site (-21) does not mediate binding. (Lower panel) Alignment of 5' flanking DNA sequences from several mammalian core histone genes showing a homologous A/T rich DNA sequence. Underlined in the diagram are sequences homologous to the CAAT-box (reviewed in 21). References: human FO108 H4 gene (11; U. Pauli, unpublished data); human ST519 H3 gene (14); human H3 gene 5B, human H3 gene 26H and mouse H3-2 gene (38). The human ST519 H3 gene is largely identical to the human H3 gene 5B.

nucleotides -119 and -76, with a demonstrated requirement for sequences located between -119 and -94.

Comparison of the DNA sequences within the H4-A and H3-A probes (shown in Figure 8) resulted in the following observations. The DNA region apparently involved in the formation of the fastest migrating DNA/protein complex (A-band) of the H4-A probe has a strong bias for purine residues in the sense(+) strand (43 purines in 50 nucleotides; see Figure 8). The H3-A DNA segment shown to be involved in the formation of a similar complex does not

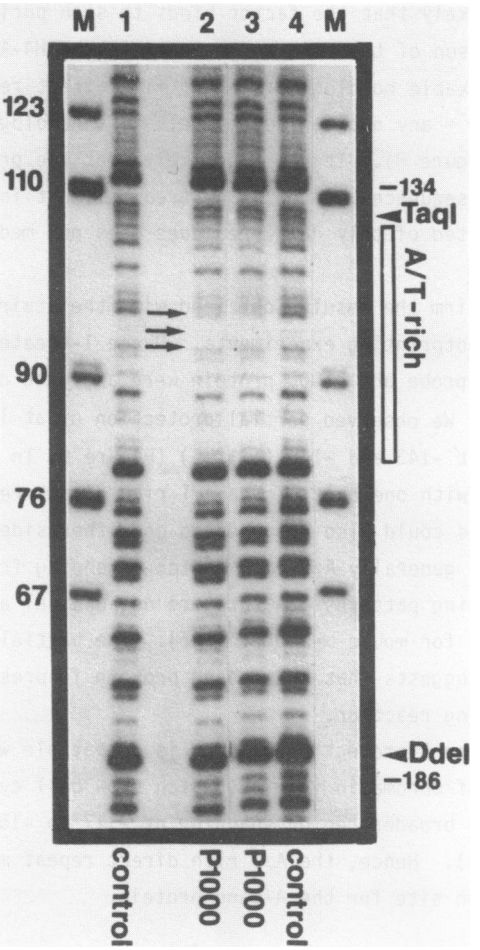


Figure 9: DNase I footprint analysis of the H4-A probe sense strand. The H4-A probe was labeled at the 5' terminal EcoRI site (nt -241), incubated with 5 µg E. coli DNA in the presence of 25 µg (ln 2) or 50 µg (ln 3) of P1000-fraction and exposed to DNase I. Ln 1 and 4: DNase I alone. Ln M: pBR322/HpaII marker (sizes are indicated by numbers at the left of the figure). Bold arrows indicate the positions of Taq I and Dde I sites in the H4-A probe. Fine arrows point to bands with decreased intensities when P1000 protein is present. Open bar indicates a generally A/T rich region.

(21 purines in 43 nucleotides). Notably, a stretch of 16 consecutive purines (at position -41 to -56) in the H3 minus-strand, partially homologous to the H4-A binding segment, can be removed without affecting the binding. Assuming that the protein involved in this DNA/protein complex is identical in both

cases, it seems unlikely that the factor binds to such purine-rich sequences. DNA sequence comparison of the binding segments in the H4-A and H3-A probes, however, shows remarkable homology of an A/T rich direct repeat in both probes (5' dATTT(N4)ATTT; N = any nucleotide); albeit this homology is found on opposite strands (see Figure 8). It seems unlikely that the protein has affinity for random A/T rich sequences, since an upstream element in the H4-A probe (nt -216 to -202) comprised of only dA/dT-residues does not mediate binding.

DNase I footprinting

In order to confirm the results obtained with the stairway assay, we have performed DNase I footprinting experiments. DNase I-treated binding mixtures containing the H4-A probe and P1000 protein were analyzed on a denaturing polyacrylamide gel. We observed partial protection of at least three sites in the region between nt -143 and -148 (± 1 nt) (Figure 9, ln 2 and 3), which exactly corresponds with one half of the A/T rich direct repeat. Slight intensity differences could also be observed on either side of the three main protected bands in a generally A/T rich region extending from nt -136 to -160. Short DNase I shielding patterns (6-9 nt) are not unusual and have, for instance, been found for mouse α-protein (36). The partial protection pattern, moreover, suggests that the A-band protein is present in limited amounts in the binding reaction.

The observed DNase I protection pattern is compatible with DNase I protection studies of chromatin *in vivo*, which show cell cycle independent protection of a much broader region spanning nt -117 to -150 (U. Pauli *et al*, manuscript submitted). Hence, the A/T rich direct repeat might constitute part of a recognition site for the A-band protein.

CONCLUSION

We have detected and partially purified a nuclear DNA binding protein with affinity for a segment within the functional promoter of a cell cycle dependent human H4 histone gene (F0108). This factor shows no affinity for a DNA fragment upstream of the H4 histone gene promoter. The protein binding segment of the H4 gene (nt -183 to -134) lacks known RNA polymerase II consensus sequences and is located in a region that has previously been shown to contain cell cycle stage specific DNAse I and S1 hypersensitive sites (13). In addition, we have partially purified a protein that binds to an upstream segment (nt -119 to -76) within the *in vivo* promoter of a cell cycle dependent H3 gene (ST519).

Both proteins appear similar, because they: (I) cause a comparable probe

retardation, with respect to relative migration rates and intensities of the shifted A-bands, in gel retardation assays; (II) are impaired in binding to putative cognate sequences in the presence of the alternating copolymer poly dI/dC, but not in the presence of E. coli DNA; (III) co-elute over both a DEAE-Sephacel and a phosphocellulose chromatography column; (IV) have comparable overall DNA binding activities at several times during the cell cycle; and (V) recognize DNA segments in both the H3 and H4 histone promoters that contain homologous A/T rich DNA sequences.

Taken together these data suggest that the two DNA binding factors are identical. Therefore, we have designated the protein Histone promoter-associated Nuclear Factor-A (HiNF-A) and propose that HiNF-A is a sequence-specific DNA binding protein with a dA/dT rich consensus sequence.

The overall DNA binding activity of HiNF-A is not dramatically altered during the cell cycle. This finding is most consistent with a possible role for HiNF-A in the basal expression of cell cycle dependent H3 and H4 histone genes, rather than in the cell cycle stage specific modulation of human histone gene expression. Further studies on the characterization of the protein, its DNA recognition sequence and possible function are in progress.

ACKNOWLEDGEMENTS

These studies were supported by grants from NIH (GM32010), NSF (PCM83-18177) and the March of Dimes Birth Defects Foundation (1-813). AJvW gratefully acknowledges the encouragement and the private financial support that enabled him to pursue this project. We thank Dr. Peter van der Vliet and Erik de Vries for advice and materials and Pauli Walton for editorial assistance. Special thanks to Dr. Urs Pauli, Dr. Susan Chrysogelos, and Dr. Harry Nick for helpful discussions and comments on the manuscript.

REFERENCES

1. Sierra, F., Lichtler, A., Marashi, F., Rickles, R., Van Dyke, T., Clark, S., Wells, J., Stein, G. and Stein, J. (1982) Proc. Natl. Acad. Sci. USA 79, 1795-1799.
2. Carozzi, N., Marashi, F., Plumb, M., Zimmerman, S., Zimmerman, A., Coles, L.S., Wells, J.R.E., Stein, G. and Stein, J. (1984) Science 224, 1115-1117.
3. Zwollo, P., Stein, G.S. and Stein, J.L. (1984) Biochem. Biophys. Res. Comm. 124, 988-993.
4. Heintz, N., Zernik, M. and Roeder, R.G. (1981) Cell 18, 661-668.
5. Green, L. , van Antwerpen, R., Stein, J., Stein, G., Triputti, P., Emanuel, B., Selden, J. and Croce, C. (1984) Science 226, 838-840.
6. Triputti, P., Emanuel, B.S., Croce, C.M., Green, L.G., Stein, G.S and Stein J.L. (1986) Proc. Natl. Acad. Sci. USA 83, 3185-3188.

7. Stein, G.S., Stein, J.L. and Marzluff, W.F. (eds.) (1984) "Histone Genes" John Wiley and Sons, New York.
8. Stein, G.S. and Stein, J.L. (1984) BioEssays 1, 202-206.
9. Schümperli, D. (1986) Cell 45, 471-472.
10. Plumb, M., Stein, J. and Stein, G. (1983) Nucl. Acids Res. 11, 2391-2410.
11. Sierra, F., Stein, G. and Stein, J. (1983) Nucl. Acids Res. 11, 7069-7086.
12. Moreno, M.L., Chrysogelos, S.A., Stein, G.S. and Stein, J.L. (1986) Biochemistry, 25, 5364-5370.
13. Chrysogelos, S., Riley, D.E., Stein, G. and Stein, J. (1985) Proc. Natl.Acad. Sci. USA 82, 7535-7539.
14. Marashi, F., Helms, S., Shiels, A., Silverstein, S., Greenspan, D., Stein, G. and Stein, J. (1986) Biochem. Cell Biol. 64, 277-289.
15. Heintz, N. and Roeder, R.G. (1984) Proc. Natl. Acad. Sci. USA 81, 2713-2717.
16. Hanly, S.M., Bleecker, G.S. and Heintz, N. (1985) Mol. Cell. Biol. 5, 380-389.
17. Shiels, A., Marashi, F., Stein, J. and Stein, G. (1985) Biochem. Biophys. Res. Comm. 127, 239-246.
18. Green, L., Schlaffer, I., Wright, K., Moreno, M.L., Berand, D., Hager, G., Stein, J. and Stein, G. (1986) Proc. Natl. Acad. Sci. USA 83, 2315-2319.
19. Capasso, O. and Heintz, N. (1985) Proc. Natl. Acad. Sci. USA 82, 5622-5626.
20. Kroeger, P., Stewart, C., van Wijnen, A.J., Shiels, A., Stein, J. and Stein, G., manuscript submitted.
21. Dynan, W.S. and Tjian, R. (1985) Nature (London) 316, 774-778.
22. Dynan, W.S. and Tjian, R. (1983) Cell 32, 669-680.
23. Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) Cell 43,439-448.
24. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) Nature (London) 319, 154-158.
25. Hromas, R. and Van Ness, B. (1986) Nucl. Acids Res. 14, 4837-4848.
26. Leegwater, P.A.J., van der Vliet, P.C., Rupp, R.A.W., Nowock, J. and Sippel, A.E. (1986) EMBO J. 5, 381-386.
27. Garner, M.M. and Revzin, A. (1981) Nucl. Acids Res. 9, 3047-3060.
28. Fried, M. and Crothers D.M. (1981) Nucl. Acids Res. 9, 6505-6525.
29. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) "Molecular Cloning: a laboratory manual", Cold Spring Harbor Laboratory.
30. Challberg, M.D. and Kelly, T.J., Jr. (1979) Proc. Natl. Acad. Sci. USA 76, 655-659.
31. Bradford, M.M. (1976) Anal. Biochem. 72, 248-254.
32. Stein, G.S. and Borun, T.W. (1972) J. Cell Biol. 52, 292-307.
33. Revzin, A., Ceglarek, J.A. and Garner, M.M. (1986) Anal. Biochem. 153, 172-177.
34. Leegwater, P.A.J., van Driel, W. and van der Vliet, P.C. (1985) EMBO J.4, 1515-1521.
35. Nagata, K., Guggenheimer, R.A., Enomoto, T., Lichy, J.H. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA 79, 6438-6442.
36. Solomon, M.J., Strauss, F. and Varshavsky, A. (1986) Proc. Natl. Acad. Sci. USA 83, 1276-1280.
37. Schneider, R., Gander, I., Muller, U., Mertz, R. and Winnacker, E.L. (1986) Nucl. Acids Res. 14, 1303-1317.
38. Wells, D.E. (1986) Nucl. Acids Res. 14, Supplement r119-149.