Volume 15 Number 5 1987 Nucleic Acids Research

Nuclear proteins from lactating mammary glands bind to the promoter of a milk protein gene

Henryk Lubon and Lothar Hennighausen*

National Institutes of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Building 10, 9N113 Bethesda, MD 20892, USA

Received September 26, 1986; Revised and Accepted January 15, 1987

ABSTRACT

The gene for the whey acidic protein (WAP) is expressed specifically in the lactating mammary glands of rodents. We present evidence that nuclear proteins from mammary epithelial cells form a multiple nucleoprotein complex with the WAP gene promoter/upstream region. As monitored by mobility shifts, nuclear proteins from lactating mammary glands and from the mammary cell line MCF-7 form four high affinity complexes with a fragment spanning the region between nucleotides -175 and -88. Nuclear proteins from liver and HeLa cells generate only three high affinity complexes. DNAaseI and ExonucleaseIII protection confirmed the binding of mammary nuclear proteins to specific sequences in the WAP gene upstream region.

This is the first report to describe the interaction of nuclear proteins from lactating mammary glands with cognate binding sites in the promoter/upstream region of a milk protein gene. The possibility of the binding sites being candidates for cis-acting regulatory elements governing the regulated expression of the WAP gene is discussed.

INTRODUCTION

Synthesis and secretion of milk proteins is specific to the lactating mammary gland. The expression of the corresponding genes is subject to complex positive and negative regulation by both steroid and peptide hormones (1-4). The whey acidic protein (WAP) is the major whey protein in the milk of mice and rats and its mRNA accounts for about 15% of the total mRNA in the lactating tissue (1,5-7). The genes for the mouse and rat WAPs have been cloned and their structures determined (8). Recently we have shown that the promoter/upstream region of the mouse WAP gene confers tissue specific and hormone dependent regulation of a marker gene in transgenomic animals (9).

The expression of a number of genes has been demonstrated in transfected tissue culture cells. The dependence on the availability of appropriate cell lines is a major limitation in the analysis of inducible and tissue specific genes. Despite considerable effort by several

© IRL Press Limited, Oxford, England.

independent investigators, regulated expression of cloned milk protein genes (8,10-14) has not been observed in transfected mammary cell lines (J.Rosen, R.Craig, L.Hall, P.Qasba, personal communications and our unpublished observations). In some cases such limitations maybe overcome by introducing genes into primary cells (15). Since dissociated mammary epithelial cells do not maintain all their differentiated functions (16), however, this approach is not readily amenable to experimental analysis of milk protein gene expression. When primary mammary epithelial cells are grown on a suitable substratum no mRNA for the whey acidic protein and α-lactalbumin can be detected (16,17). Bissell and coworkers (17) correlated the absence of regulated milk protein gene expression with the lack of 'factors', matrix, extracellular stimuli and/or proper cell-cell interactions.

The search for sequences in the promoter region of genes that bind nuclear proteins with high affinity is an alternative approach to locate regulatory elements. Direct _in vitro_ interactions of nuclear proteins with sequences in potential regulatory elements has been described to our knowledge in thirteen polymerase II genes from higher eukaryotes (15,18-37). However, in many of these cases nuclear proteins were derived from heterologous cells which do not express the gene under investigation. A correlation between protein binding and transcriptional activation has been made in certain cases where _in vivo_ and _in vitro_ expression systems are available (24,26,33). In an attempt to define potential control elements within the functional WAP promoter/upstream region (8,9), we searched for sequences that are specifically recognized by nuclear proteins from lactating mammary glands, from a mammary epithelial cell line and from non mammary cells. The identification of sequences recognized by mammary gland nuclear proteins can be considered as a preselection for elements governing WAP gene expression, which thereafter can be tested using a transgenomic system.

MATERIALS AND METHODS

Cell culture

MCF-7 cells were propagated in RPMI medium supplemented with penicillin and streptomycin and 10% fetal calf serum (Gibco).

Preparation of nuclei and nuclear extracts

Confluent MCF-7 monolayers were washed with cold Dulbecco's phosphate buffered saline, scrapped with a rubber policeman into a small volume of

the same buffer and centrifuged at 800xg for 5 minutes. The cell pellet was resuspended and washed once in this buffer. Pelleted cells were resuspended in 10 volumes of phosphate buffered saline at 4°C and centrifuged as above. Nuclei were prepared from pelleted cells as follows: cells were resuspended in 5 pellet volumes of 0.3M sucrose in buffer A (10mM HEPES at pH 7.9, 10mM KCl, 1.5mM $MgCl_2$, 0.1mM EGTA, 0.5mM dithiothreitol, 0.5mM phenylmethylsulfonyl fluoride and 2 ug/ml each of antipain, leupeptin and pepstatin A). Cells were lysed by 12 strokes with a B pestle in a Dounce glass homogenizer and by an additional 1 to 2 strokes in the presence of 0.3 to 0.4% NP-40. Completion of lysis was monitored using a phase contrast microscope. The homogenate was centrifuged at 1200xg for 10 minutes and the pelleted nuclei were washed twice in 0.3M sucrose in buffer A without NP-40.

Mammary glands and livers were dissected from five 12-day lactating rats and stored at -70°C. Nuclei were isolated by pulverizing the frozen tissue followed by a homogenization step in 10 volumes of 0.3M sucrose in buffer A using a glass-teflon homogenizer (1500 rpm, 12 strokes) and 2 to 3 additional strokes in the presence of 0.5% NP-40. The optimal number of strokes was determined by microscopy. The homogenate was filtered through nylon mesh and crude nuclei were pelleted by centrifugation at 1200xg for 10 minutes. The pellet was resuspended in 0.3M sucrose, layered onto 1.7M sucrose in buffer A and centrifuged in the HB4 rotor at 12500 rpm for 1 hour. After washing the nuclei once in 0.3M sucrose in buffer A, they were recentrifuged at 1200xg for 10 minutes. Each preparation of nuclei was used immediately for extraction of protein. Nuclear extracts from mammary glands and MCF-7 cells were prepared using a modification of the method of Dignam _et al_. (38) according to the procedure described by Wu (27). Crude nuclear extract from MCF-7 cells and lactating mammary glands was fractionated on a heparin-agarose column (27) and stepped washed with 0.1M, 0.3M and 0.6M NaCl.

Recombinant plasmids and isolation of DNA fragments

The 378 bp SstI-KpnI (-354 to +24) promoter/upstream fragment, the 266 bp SstI-XbaI (-354 to -88) upstream fragment and the 112 bp XbaI-KpnI (-88 to +24) promoter fragment from the mouse WAP gene (8) and the 207 bp Sau3A-BamHI (-202 to +5) fragment from the rat α-lactalbumin gene (13) were cloned into the polylinker of pUC18. The plasmid harboring a fragment overlapping the chicken lysozyme gene promoter from -208 to +15 (15) was a gift from Gunther Schutz. DNA fragments were isolated from low-melting

agarose gels. Restriction fragments with 5' overlapping ends were labeled either by 'fill in' using ^{32}P α-dNTPs and the Klenow enzyme or with ^{32}P γ-ATP and polynucleotide kinase. Blunt end fragments were labeled with ^{32}P γ-ATP.

DNA-protein mobility shift assay

Mobility shift assays were performed as described (39,40). The labeled fragments, sonicated *E. coli* DNA or poly(dI-dC) (Pharmacia) and nuclear extract were mixed in 25ul of 0.1% Triton X-100, 12% glycerol, 1mM Na-EDTA, 5mM $MgCl_2$, 1mM DTT, 10mM Tris-HCl at pH 7.5 and 80mM NaCl. One ug of nuclear extract was added last. The mixture was incubated for 40 to 60 minutes at room temperature. Samples were loaded onto a low-ionic-strength 4% polyacrylamide gel (acrylamide:bisacrylamide=30:1) containing 6.6mM Tris-HCl at pH 7.5, 1mM EDTA and 3.3mM Na-acetate. The gel (0.15x16 cm) was preelectrophoresed for 2 hours at 20 mA. The buffer was recirculated between compartments. Electrophoresis was carried out at 30 mA and room temperature until the bromophenol blue had run to the bottom of the gel. After soaking the gel in 5% glycerol, 5% methanol and 5% acetic acid, it was transferred to Whatman 3MM paper, dried and subjected to autoradiography.

ExonucleaseIII protection analysis

The DraI-XbaI restriction fragment from the WAP gene upstream region (Fig. 1) was labeled at either 5' end. One ng the fragments were incubated with 5ug poly(dI-dC) and 1ug mixed deoxynucleotides $p(dN)_5$ (Pharmacia) in the absence or presence of 1ug nuclear proteins, in 25ul binding buffer (0.1% Triton X-100, 5% glycerol, 0.1mM EDTA, 5mM $MgCl_2$, 1mM DTT, 10mM Tris-HCl at pH 7.5, 80mM NaCl). After incubation at room temperature for 30 minutes, ExonucleaseIII (Bethesda Research Laboratories) was added and incubation lasted for another 10 minutes at 30°C. The reaction was terminated by the addition of an equal volume of 20mM EDTA and 1% SDS. The reaction products were analysed after purification on a 10% polyacrylamide/8M urea sequencing gel and subjected to autoradiography.

DNAaseI protection analysis

The SstI-KpnI (-354 tp +24) fragment was cloned into the polylinker of pUC18. Upon cutting with BamHI (located within the polylinker next to the KpnI site) the lower strand was labeled with ^{32}P γ-ATP (5' label) and the upper strand with ^{32}P α-dGTP (3' label). The labeled fragments were recut with SstI and eluted from an agarose gel. One ng of the labeled probe was incubated with 1 ug poly(dI-dC) and increasing amounts of

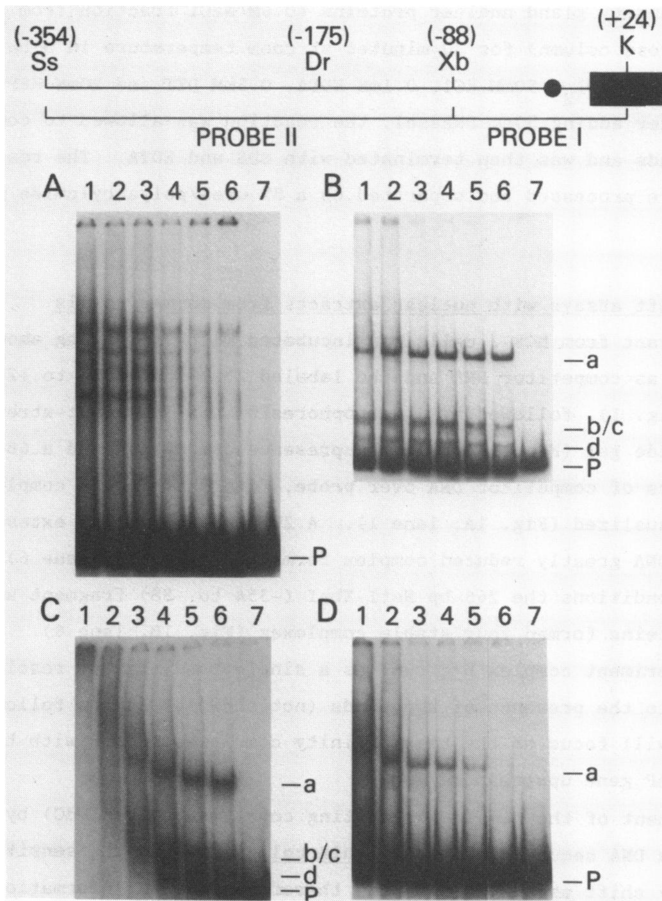


Figure 1
Mobility shift complexes formed between mammary cell derived nuclear proteins and the WAP promoter. In the top part the map of the promoter is shown. The solid box indicates the first exon and the solid circle the 'Goldberg-Hogness' (TTTAAA) box. Dr=DraI, K=KpnI, Ss=SstI, Xb=XbaI. The numbers indicate the positions of the restriction sites with respect to the cap site. P indicates the positions of the naked DNA probes and a-d represent protein-DNA complexes. Panel A) 0.1 ng of the labeled 112 bp XbaI-KpnI fragment (probe I) was incubated with 1 ug protein from the 0.3M NaCl extract of MCF-7 nuclei and with increasing amounts of poly(dI-dC) competitor DNA, followed by electrophoresis in a low-ionic-strength 4% polyacrylamide gel. Lanes 1-6) probe with extract and 680, 1360, 2720, 5440, 10880 and 21760-fold weight excess of poly(dI-dC) per assay respectively; lane 7) probe. Panel B) As panel A, but with 0.1 ng of the labeled 266 bp SstI-XbaI fragment (probe II). Panel C) As panel B, but instead of MCF-7 extract nuclear extract from lactating rat mammary glands was assayed. Panel D) As panel B, but the poly(dI-dC) was replaced by unlabeled and sonicated *E. coli* competitor DNA.

lactating mammary gland nuclear proteins (0.6M NaCl fraction from the heparin-agarose column) for 30 minutes at room temperature in a buffer containing 1mM $MgCl_2$, 50mM KCl, 0.1mM EDTA, 0.5mM DTT and 20mM HEPES at pH 7.5. After adding 50ng DNAaseI, the reaction was allowed to continue for 30 seconds and was then terminated with SDS and EDTA. The reaction products were processed and separated on a 8% urea/polyacrylamide gel.

RESULTS

Mobility shift assays with nuclear extracts from mammary cells

Nuclear extract from MCF-7 cells was incubated with increasing amounts of poly(dl-dC) as competitor DNA and the labeled Xbal-KpnI (-88 to +24) fragment (Fig. 1), followed by electrophoresis in a low-ionic-strength 4% polyacrylamide gel (Fig. 1A). In the presence of protein and a 680-fold weight excess of competitor DNA over probe, five protein-DNA complexes could be visualized (Fig. 1A, lane 1). A 20,000-fold weight excess of competitor DNA greatly reduced complex formation (Fig. 1A, lane 6). Under identical conditions the 266 bp SstI-XbaI (-354 to -88) fragment and MCF-7 nuclear proteins formed four stable complexes (Fig. 1B, lane 6). Although in this experiment complex b/c runs as a single band, higher resolution gels indicate the presence of two bands (not shown). In the following studies we will focus on the high affinity complexes formed with the SstI-XbaI WAP gene upstream region.

Replacement of the simple alternating copolymer poly(dI-dC) by the heterologous DNA sequence of *Escherichia coli* decreases the sensitivity of the mobility shift assay (20) and can therefore provide information about the stability of protein-DNA complexes. We repeated the experiment shown in Figure 1B with *E. coli* DNA as competitor. Only complex a was formed (Fig. 1D) suggesting that its protein component is either very abundant or has a high affinity for its cognate binding site.

MCF-7 cells are of mammary origin but they do not express endogenous milk protein genes. As shown in other systems, however, this does not necessarily mean that specififc regulatory and DNA-binding proteins are absent (24). We therefore repeated the binding experiments on the SstI-XbaI fragment with nuclear extract from lactating mammary glands, the tissue in which the WAP gene is expressed physiologically. The pattern of complexes obtained with lactating mammary gland extract matched that obtained with MCF-7 extract, but complex d is more abundant in the physiologically active cells. (Fig. 1C).

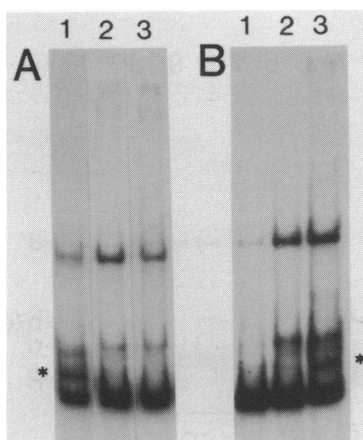


Figure 2
Protein-DNA complexes formed with nuclear extracts from lactating rat mammary glands, rat livers, HeLa cells and MCF-7 cells. A) 0.1 ng of the labeled SstI-XbaI fragment and 1 ug poly(dI-dC) were incubated with 1 ug crude nuclear proteins from lactating mammary glands (lane 1), rat liver (lane 2) and HeLa cells (lane 3) followed by gel electrophoresis. B) MCF-7 nuclear extract was passed over a heparin-agarose column and different fractions were tested with the SstI-XbaI fragment in the mobility shift assay. Lanes 1 to 3: 100 ng protein from the 0.1M NaCl, 0.3M NaCl and 0.6M NaCl fraction were incubated with 1 ug poly(dI-dC) and 0.1 ng probe followed by gel electrophoresis. * indicate complex d.

Mobility shift assays with nuclear extracts from non mammary cells

Upon incubation of the SstI-XbaI fragment with nuclear extracts from rat liver or HeLa cells complexes a and b/c were generated, but the formation of complex d was at least greatly reduced (Fig. 2A, lanes 2 and 3). Incubation of the SstI-XbaI fragment with a mixture of HeLa extract and either MCF-7 extract or mammary gland extract resulted in the formation of complex d (data not shown). This eliminates the possibility that proteases in the non mammary extracts inactivate the protein component of complex d.

To determine whether the four complexes were generated by distinct protein components of the nuclear extract, the MCF-7 nuclear proteins were chromatographed on heparin-agarose and the fractions were tested in the DNA binding assay (Fig. 2B). While the 0.3M NaCl fraction generated mainly complexes a and b/c, the 0.6M NaCl fraction generated all four complexes. This suggests that the protein component of complex d is distinct from the protein components of complexes a and b/c.

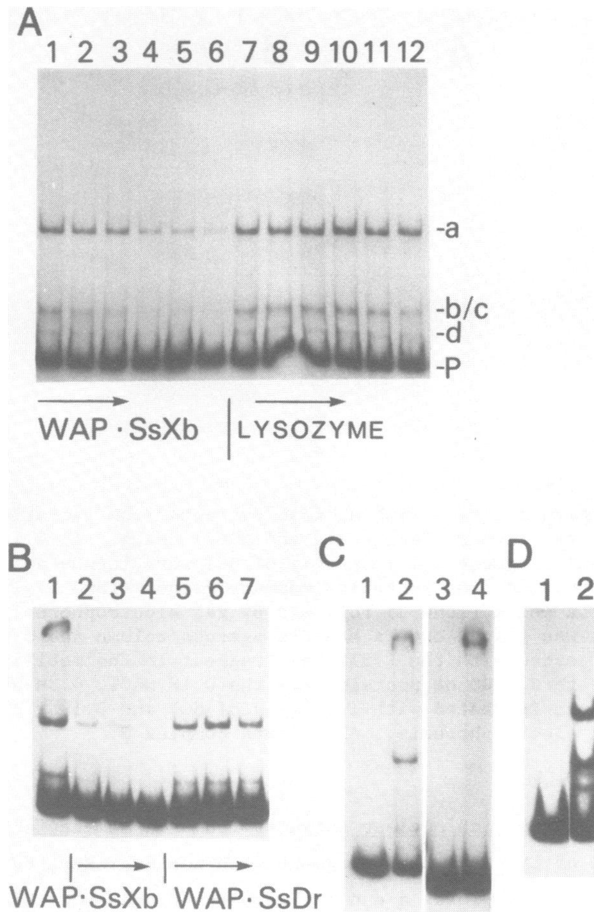


Figure 3
Binding and competition analysis in MCF-7 nuclear extracts. One ng of the SstI-XbaI probe (Fig. 1) and 10 ug poly(dI-dC) were incubated with increasing amounts of competitor DNA and processed as described in Fig. 1. Panel A: lanes 1-6) 0, 6.25, 12.5, 25, 50 and 100 ng of the unlabeled SstI-XbaI fragment were added; lanes 7-12) 0, 6.25, 12.5, 25, 50 and 100 ng of a 250 bp fragment overlapping the chicken lysozyme promoter (-208 to +15) were added. Panel B: lane 1) complex formation in the absence of additional cloned competitor DNA, lanes 2-4 and 5-7 represent assays with 50, 100 and 150 ng of the SstI-XbaI and the 180 bp SstI-DraI fragment respectively. Due to the cloning strategy the SstI-DraI piece is flanked by a polylinker, adding an additional 45 bp. Panel C: complexes formed with the 265 bp SstI-XbaI fragment (lane 2) and the 185 bp SstI-DraI fragment (+45 bp polylinker) (lane 4). Lanes 1 and 3 show the probe. Panel D: naked 87 bp DraI-XbaI fragment (lane 1) and complexes formed between the DraI-XbaI fragment and crude MCF-7 extract (lane 2).

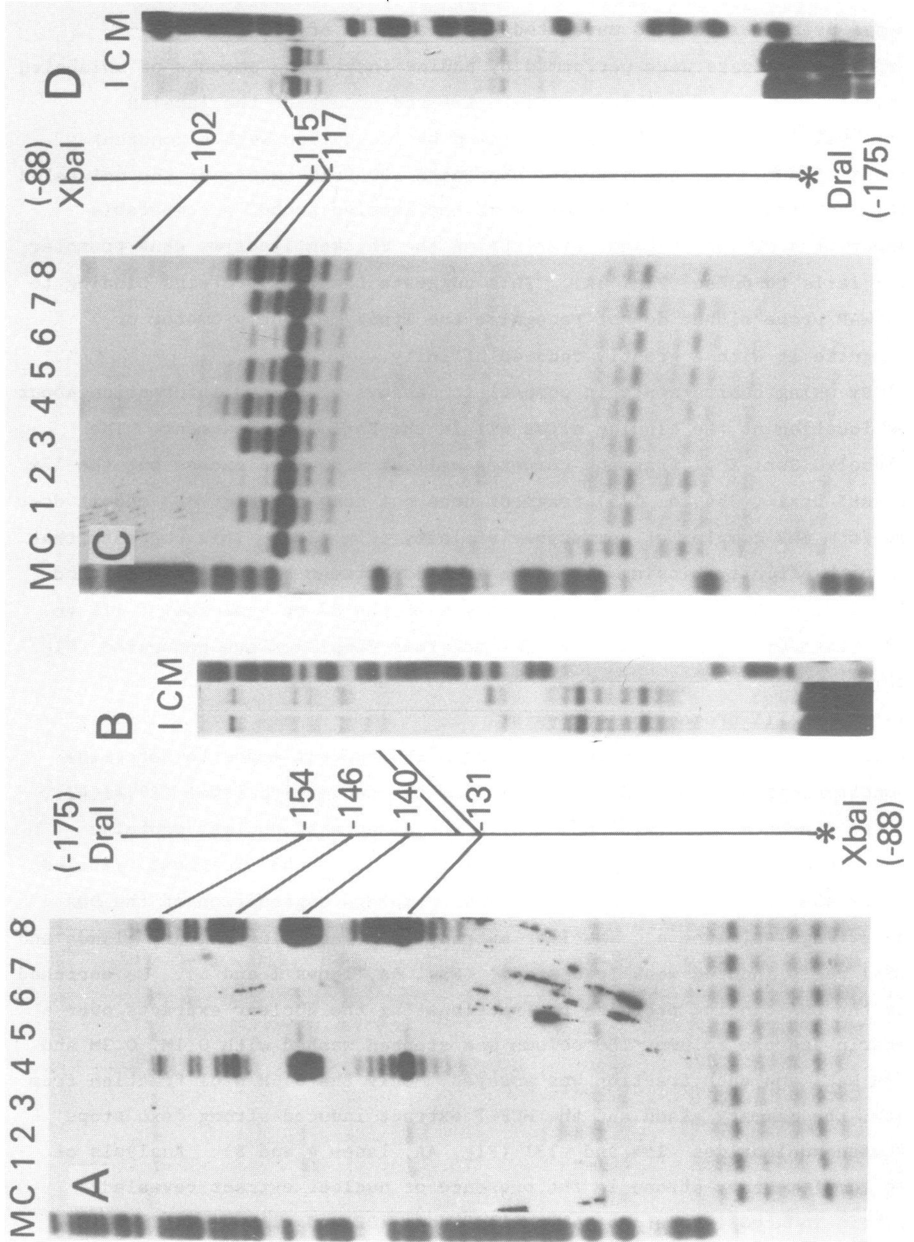
Sequence specificity of the protein-DNA complexes

Specificity of the protein-DNA complexes was determined in competition assays using related and unrelated DNA fragments of similar size. Competition assays were performed by adding increasing amounts of unlabeled specific competitor DNA fragments to a fixed amount of the labeled SstI-XbaI fragment (probe II), followed by incubation with a constant amount of MCF-7 nuclear extract. Whereas a 100-fold excess of the unlabeled SstI-XbaI fragment competes well with the labeled probe, a comparable amount of a 250 bp fragment overlapping the chicken lysozyme gene promoter (15) fails to do so (Fig. 3A). This suggests that the proteins binding to the WAP probe either do not recognize the lysozyme gene promoter or recognize it with a greatly reduced affinity.

By using subfragments in competition assays we gained information about the location of the binding sites within the SstI-XbaI fragment. The unlabeled Sstl-XbaI fragment competes well at a 50-fold excess but the 179 bp SstI-DraI (-354 to -175) fragment does not compete (Fig. 3B) and it does not form any complex of comparable affinity (Fig. 3C). This suggests that the high affinity binding sites are located between nucleotides -175 and -88. Indeed, in a mobility shift assay with the 87 bp DraI-XbaI (-175 to -88) fragment and MCF-7 nuclear extract four complexes are generated (Fig. 3D).

ExonucleaseIII protection analysis

ExonucleaseIII protection analysis is an independent means to determine boundaries of protein-DNA interaction at the nucleotide level (27,41,42). In the presence of poly(dI-dC) and in the absence of nuclear proteins ExonucleaseIII digestion of the DraI-XbaI (-175 to -88) fragment (labeled at the XbaI site) resulted in an almost complete degradation of the DNA probe (Fig. 4A, lane c). Nuclear extract from lactating mammary glands and MCF-7 cells induced weak 'exo stops' (Fig. 4A, lanes 1 and 5). We enriched for the DNA binding proteins by fractionating the nuclear extracts over a heparin-agarose column. The column was stepped washed with 0.1M, 0.3M and 0.6M NaCl and each fraction was assayed. Only the 0.6M NaCl fraction from both, the mammary gland and the MCF-7 extract induced strong 'exo stops' between nucleotides -154 and -131 (Fig. 4A, lanes 4 and 8). Analysis of the complementary strand in the presence of nuclear extract revealed protein induced 'exo stops' at nucleotides -102 and -115/-117 (Fig. 4C, lanes 1 and 5). Proteins in the 0.6M NaCl fraction enhanced these stops and induced an additional one at nucleotide -113 (Fig. 4B, lanes 4 and 8).


A
MC 1 2 3 4 5 6 7 8
(-175)
Dral
-154
-146
-140
-131
Xbal
(-88)
B
I C M
C
MC 1 2 3 4 5 6 7 8
(-88)
Xbal
-102
-115
-117
Dral
(-175)
D
I C M

Proteins in the 0.3M NaCl fraction induced only weak stops on the lower strand (XbaI label), but strong stops on the upper strand (DraI label). This is reproducible and could be interpreted as preferential binding to one strand.

As predicted from the mobility shift assays, a less complex ExonucleaseIII protection pattern was obseved in the presence of E. coli competitor DNA (Fig. 4B and 3D). The relatively weak protein induced 'exo stops' at nucleotides -133/-131 (XbaI label, Fig. 4B, lane 1) and nucleotide -117 (DraI label, Fig. 4D, lane) match those seen in the presence of poly(dI-dC), thereby confirming the presence of a protected element. The 'exo stops' observed on the naked DNA fragments are the result of incomplete digestion, suggesting that E. coli DNA is a better substrate for ExonucleaseIII than is poly(dI-dC). Additional weak 'exo stops' in the presence of nuclear proteins and poly(dI-dC) are probably caused by the stuttering of ExonucleaseIII in the immediate vicinity of the protein binding sites.

Competition analysis

Specificity of the protein induced 'exo stops' was verified in a protection-competition analysis. The DraI-XbaI fragment labeled at either end, poly(dI-dC) and competitor DNA fragments were mixed prior to incubation with the 0.6M NaCl heparin-agarose fraction from the MCF-7 extract, followed by ExonucleaseIII digestion. A 25-fold weight excess of the unlabeled WAP gene SstI-XbaI fragment over the probe (labeled at the XbaI site) resulted in the disappearance of all major 'exo stops' (Fig. 5A,

Figure 4

ExonucleaseIII protection analysis of the DraI-XbaI (-175 to -88) fragment. One ng of the XbaI-DraI fragment, labeled at either the XbaI (panels A and B) or DraI (panels C and D) site, together with 5 ug of poly(dI-dC) (panels A and C) or 1.5 ug E. coli DNA (panels B and D), were digested with 520 units ExonucleaseIII in the absence or presence of nuclear proteins. The reaction products were separated on a 10% polyacrylamide/8 M urea gel and subjected to autoradiography. Panel A: M) A/G Maxam and Gilbert sequencing reaction, lane C) digestion products of the probe in the absence of nuclear proteins. Lanes 1-4 and 5-8 show the ExonucleaseIII digestion products in the presence of nuclear proteins from MCF-7 cells and lactating mammary glands, respectively. Lanes 1 and 5) crude nuclear proteins, lanes 2 and 6) 0.1M NaCl fraction from the heparin-agarose column, lanes 3 and 7) 0.3M NaCl fraction, lanes 4 and 8) 0.6M NaCl fractions. Panel B: M) A/G Maxam and Gilbert sequencing reaction, lane C) digestion products of the probe in the absence of nuclear proteins, lane 1) digestion products in the presence of the 0.6M NaCl fraction from MCF-7 nuclear proteins. Panel C and panel D match panel A and panel B respectively, with the exception that the fragment was labeled at the DraI site.

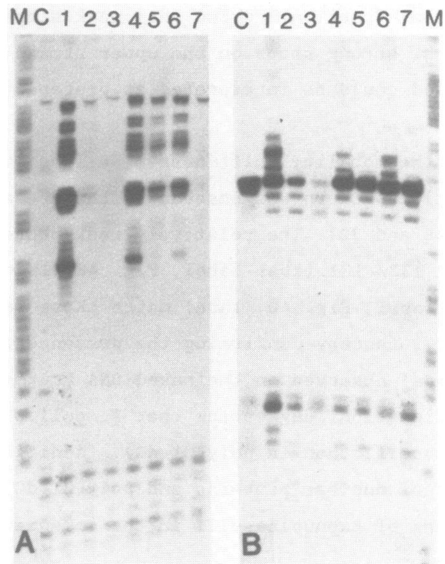


Figure 5
ExonucleaseIII protection and competition analysis. Panel A: One ng of the DraI-XbaI fragment labeled at the XbaI site and 5 ug poly(dI-dC) were digested with 520 units ExonucleaseIII in the absence (lane C) or presence (lane 1) of 1 ug of the 0.6M NaCl heparin-agarose fraction from MCF-7 nuclear protein. Lanes 2-7 show the reaction products of the competition assays in the presence of: 2.5 and 10 nanogram of the SstI-XbaI WAP gene fragment (lanes 2 and 3), 2.5 and 10 nanogram of the 250bp lysozyme gene promoter fragment (lanes 4 and 5) and 2.5 and 10 nanogram of the 207bp Sau3A-BamHI fragment overlapping the rat α-lactalbumin gene promoter (lanes 6 and 7). Panel B: the DraT-XbaI fragment was labeled at the DraI site; otherwise all lanes are comparable with their counterparts in panel A. Lane M) A/G Maxam and Gilbert sequencing reactions of the respective fragments.

lane 2). A 250 bp fragment spanning the chicken lysozyme gene promoter acted as a poor competitor (Fig. 5A, lanes 4 and 5) and a 200 bp fragment overlapping the α-lactalbumin gene promoter and upstream region competed specifically but less well than the WAP gene fragment (Fig. 5A, lanes 6 and 7). Competition analysis on the upper strand of the DraI-XbaI fragment with the SstI-XbaI fragment as competitor DNA gave complementary results (Fig. 5B, lanes 2 and 3). With the amount of competitor used, both the α-lactalbumin gene fragment and the lysozyme gene fragment competed equally well for binding observed on the upper strand (Fig. 5B, lanes 4-7). Competition for protection on the upper strand is more sensitive to competitor DNA which explains the results shown in Fig. 5B.

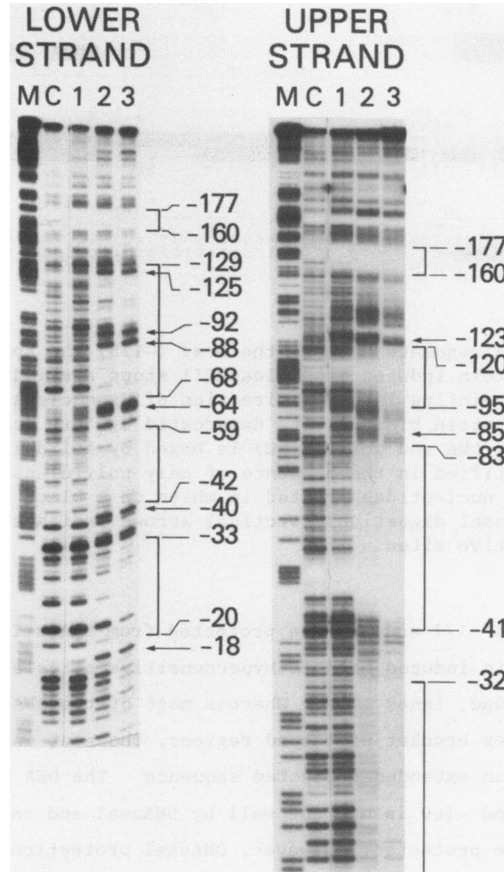


Figure 6
DNAaseI protection analysis of the SstI-KpnI fragment. M: A/G Maxam and Gilbert sequencing reaction of the respective fragments; C: DNAaseI digestion products of the DNA fragments in the absence of nuclear proteins; lanes 1-3: DNAaseI digestion products of fragments incubated with 6ug, 12ug and 24ug nuclear proteins from lactating mammary glands. Protected regions are indicated by brackets and the arrows point to protein induced DNAaseI hypersensitive sites. The numbers refer to nucleotide positions with respect to the cap site.

DNAaseI protection analysis

The SstI-KpnI (-354 to +24) fragment was labeled at the KpnI site and incubated with 6ug, 12ug and 24ug of the 0.6M NaCl heparin-agarose fraction from the mammary gland extract, follwed by DNAaseI digestion and separation of the reaction products on a urea/polyacrylamide gel. Four regions

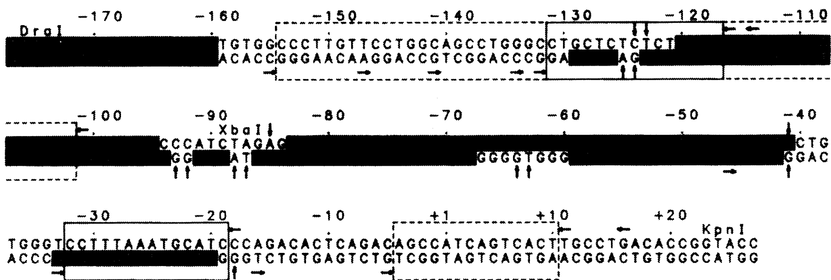


Figure 7
The double stranded sequence between the DraI (-175) and KpnI (+24) sites is presented. Protein induced ExonucleaseIII stops are indicated by horizontal arrows pointing into the direction of ExonucleaseIII digestion (3' to 5'). The protein binding site demarcated by 'exo stops' in the presence of *E. coli* DNA and poly(dI-dC) is boxed by solid lines. The binding sites identified in the presence of only poly(dI-dC) are boxed by dashed lines. The nucleotides printed in white on a black background are protected from DNAaseI digestion. Vertical arrows indicate protein induced DNAaseI hypersensitive sites.

between nucleotides -177 and -20 are protected from DNAaseI digestion and at least six protein induced DNAaseI hypersensitive sites are generated (Fig. 6, lower strand, lanes 1-3). Whereas most of the DNAaseI hypersensitive sites bracket protected regions, those at nucleotides -92 and -87 interrupt an extended protected sequence. The DNA between nucleotides -177 and -160 is not cut well by DNAaseI and only the central two bp appear to be protected. However, DNAaseI protection analysis on subcloned fragments revealed that the entire sequence between -177 and -160 is protected by nuclear proteins from lactating mammary glands (data not shown). Although the protection pattern observed on the upper strand is not as clear as that on the lower strand, the data are complementary (Fig. 6, upper strand, lanes 1-3). In analogy to the situation on the lower strand, a protein induced DNAaseI hypersensitive site at nucleotide -85 separates the protected region spanning the sequence between nucleotides -120 and -41. Because of unequal cutting by DNAaseI, the protected 'TTTAAA' motif is not clearly visible and not marked in Figure 7. No protection was observed upstream of nucleotide -177.

DISCUSSION

Based on mobility shift, DNAaseI and ExonucleaseIII protection assays, we report for the first time interactions of multiple nuclear proteins from

lactating mammary glands with promoter/upstream sequences of a milk protein gene. The lack of cell culture systems mimicking the physiology of lactating mammary glands made the search for regulatory elements governing milk protein gene expression a frustrating and unsuccessful one. Therefore, the <u>in vivo</u> significance of the binding sites remains to be determined.

The XbaI-KpnI (-88 to +24) promoter fragment of the WAP gene and nuclear proteins from the mammary cell line MCF-7 formed five complexes in the mobility shift assay. One complex is generated by a protein(s) recognizing the 'TTTAAA' motif, and two complexes are associated with sequences upstream and downstream of the 'TTTAAA' motif (data not shown). Four high affinity complexes were formed between the SstI-XbaI (-354 to -88) fragment and nuclear proteins from either MCF-7 cells or lactating mammary glands. Complex <u>d</u> is more pronounced in the extract from the lactating tissue, indicating an enrichment of its protein component in the physiologically active cell. Upon incubation of the SstI-XbaI fragment with nuclear proteins from liver or HeLa cells complex <u>d</u> is greatly reduced or absent, suggesting its mammary specificity.

Protein protected sequences in the WAP gene promoter/upstream region were defined independently by ExonucleaseIII protection and competition experiments. The protein-DNA complexes formed in nuclear extract were detectable in the mobility shift assay, but not very clear with ExonucleaseIII. Clear ExonucleaseIII protection was observed after enriching the nuclear DNA binding proteins from lactating mammary glands and MCF-7 cells by heparin-agarose chromatography. In the presence of <u>E. coli</u> DNA as competitor, proteins in the 0.6M NaCl heparin-agarose fraction form one complex between nucleotides -133/-131 and -117/-115. Under comparable conditions only one high affinity complex was seen in the mobility shift assay and its borders can now be defined as nucleotides -133/-131 and -117/-115. In the presence of poly(dI-dC) additional protein induced 'exo stops' demarcate the sequences between nucleotides -154 and -104, but a correlation with the mobility shift bands <u>b/c</u> and <u>d</u> has not been made. Although complex <u>d</u> is preferentially formed in mammary extract, we have not been successful in detecting differences in the 'exo stop' pattern between nuclear extracts from mammary cells and HeLa cells (data not shown). This could be explained by the sensitivity of the ExonucleaseIII assay versus the mobility shift assay, or by the possibility that the protected sequences in complex <u>d</u> are shared with protected

sequences in one of the other complexes, a precedent of which has been described recently (20).

As a third criteria for defining protein protected sequences we have performed DNAaseI footprint analyses. With the 0.6M NaCl heparin-agarose fraction from mammary nuclear extract several protected areas were detected on the lower strand of the SstI-KpnI fragment. The protected regions from nucleotides -59 to -42 and -33 to -20 encompass the sequences CCAAAGT and TTTAAA which resemble promoter elements. The observed binding is therefore presumably due to proteins recognizing these sequences. The sequence between nucleotides -177 and -160 is protected from DNAaseI cutting, but not from ExonucleaseIII digestion. This can be explained by the nature of fragments used in the two sets of experiments. Whereas the fragment used in the DNAaseI experiment contained the DraI site and WAP gene 5' flanking sequences, the fragment used in the ExonucleaseIII assay did contain only half of the DraI site and no WAP gene 5' flanking sequences. In this fragment part of the binding site was deleted and probably non-functional. Protein induced DNAaseI hypersensitive sites are indicators for protein-DNA interactions in their proximity. Two of the hypersensitive sites (-92 and -87) are located within the protected region between nucleotides -129 and -68, indicating the binding of two independent proteins to this sequence. This interpretation is supported by mobility shift assays and ExonucleaseIII protection analyses using subfragments spanning different parts of this region (data not shown). Several 'exo stops' but no DNAaseI protection were observed on the lower strand between nucleotides -154 and -132. This can be explained by the high sensitivity of ExonucleaseIII protection assay but binding to these sequences needs to be confirmed with synthetic oligonucleotides.

In order to identify protected sequences within complexes visualized in the mobility shift assay we attempted to isolate individual protein-DNA complexes after they had been treated with DNAaseI. Upon DNAaseI treatment complexes b/c and d migrated with the probe (data not shown), a problem also encountered by others (18). The weakly protected sequence in complex a spanned nucleotides -139 to -113, thereby confirming protection of this region (data not shown).

Protected Sequence Elements

Recent work by several groups has shown that genes subject to very different control mechanisms (e.g. immunoglobulin and histone genes) can be recognized by similar or identical DNA binding proteins (20,22). Therefore

the question has to be raised as to how gene- and tissue-specific the described complexes are. We compared the protected regions with other known sites of protein-DNA interaction. A 9 bp sequenece (GTGTGGCCA) between nucleotides -120 and -112 is homologous to elements in the chicken ß-globin gene upstream region (GTGTGGCCA) (32) and the Adenovirus2 major late promoter (GTGGCC) (44,45). This sequence in the chicken ß-globin gene does not seem to be protected by nuclear proteins, but is flanked by protein protected elements (32). The AD2MLP element on the other hand is protected by the AD2MLP upstream factor, a protein that also interacts weakly with the rabbit ß-globin promoter gene upstream sequence (44). A sequence matching the glucocorticoid receptor binding site consensus sequence (46) was found between -150 and -144, but its physiological significance remains to be determined.

It is conceivable that milk protein genes share one or more regulatory elements conferring mammary specific and hormone controlled expression. The sequence TGGCAGCCTGG between nucleotides -144 and -134 matches the sequence TGGCAGGATGG in a similar position in the rat α-lactalbumin gene (13). As in the WAP gene, this sequence is also bracketed by 'exo stops' in the α-lactalbumin gene (data not shown). Further sequence comparison between the promoters of the WAP gene and several other milk protein genes revealed an 11 base pair sequence with significant homologies. In the WAP gene this sequence (AAGAAGGAAGT) is located between nucleotides -112 and -103. Related sequences in the promoter regions the rat α-lactalbumin gene (TGGAGGGAAGT; 73% conservation) and the rat α-casein gene (AAGAAGGAATT; 91% conservation) are also protected by mammary gland derived nuclear proteins (data not shown). Based on this sequence homology, the protein binding data and the competition assays we suggest that the α-lactalbumin gene promoter shares protein binding sites with the WAP gene upstream region. However, this sequence is not necessarily the target for a mammary specific DNA binding protein. For example the immunoglobulin gene promoters share an octamer sequence required for B-cell specific expression which is recognized by a protein that is not B-cell specific (20). The same octameric sequence in the human histone H2B gene binds the same or a similar protein and confers a histone gene specific regulation (22). It is conceivable that one or more of the proteins bound to the WAP upstream region are members of a family of DNA binding proteins that recognize and regulate a set of promoters. Data on the significance of the

high affinity protein binding sites in the framework of mammary specific gene expression have to come from studies with transgenomic animals.

*To whom correspondence should be addressed

ACKNOWLEDGEMENTS

We thank Richard Miksicek and Gunther Schutz for the plasmid containing the chicken lysozyme gene promoter, Pradman Qasba for the rat α-lactalbumin gene promoter, Jeff Rosen for the rat α-casein gene promoter and John Clark for the ovine ß-lactoglobulin gene promoter. Roger Craig made unpublished sequence data available. The criticisms of Peter Ghazal, Uli Siebenlist, Tom Kelly, Bob Simpson and the two anonymous referees are appreciated. We are grateful to Flora Granthan for preparing mammary gland tissue. Jadwiga Nitiendel, Priscilla Furth and Mathilde Hennighausen kept up our spirits.

REFERENCES

1. Hobbs,A.A., Richard,D.A., Kessler,D.T. and Rosen,J.M. (1982) J. Biol. Chem. 257, 3598-3605.
2. Guyette,W.A., Matusik,R.J. and Rosen,J.M. (1979) Cell 17, 1013-1023.
3. Rosen,J.M., O'Neal,D.L., McHugh,J.E. and Comstock,J.P. (1978) Biochemistry 17, 290-297.
4. Topper,Y.J. (1970) Recent Prog. Horm. Res. 26, 286-308.
5. Hennighausen,L.G., Sippel,A.E., Hobbs,A.A. and Rosen,J.M. (1982) Nucleic Acids Res. 10, 3722-3742.
6. Hennighausen,L.G. and Sippel,A.E. (1982) Nucleic Acids Res. 10, 2677-2684.
7. Richard,D.A., Rodgers,J.R., Supowit,S.C. and Rosen,J.M. (1981) J. Biol. Chem. 256, 526-532.
8. Campbell,S.M., Rosen,J.M., Hennighausen,L., Strech-Jurk,U. and Sippel,A.E. (1984) Nucleic Acids Res. 12, 8685-8697.
9. Andres,A.-C., Schonenberger,C.-A., Groner,B., Hennighausen,L. LeMeur,M. and Gerlinger,P. Proc. Natl. Acad. Sci. U.S.A., in press
10. Yu-Lee,L.-Y., Richter-Mann,L., Couch,C.H., Stewart,F.A., Mackinlay,A.G. and Rosen,J.M. (1986) Nucleic Acids Res. 14, 1883-1902.
11. Jones,W.K., Yu-Lee,L.-Y., Clift,S.M., Brown,T.L. and Rosen,J.M. (1985) J. Biol. Chem. 260, 7042-7050.
12. Yu-Lee,L.-Y. and Rosen,J.M. (1983) J. Biol. Chem. 258, 10794-10804.
13. Qasba,P.K. and Safaya,S.K. (1984) Nature 308, 377-380.
14. Hall,L., Emery,D.C., Davies,M.S., Parker,D. and Craig,R.K. submitted.
15. Renkawitz,R., Schutz,G, von der Ahe,D. and Beato,M. (1984) Cell 37, 503-510.
16. Lee,E.Y.-H.P., Lee,W.-H., Kaetzel,C.S., Parry,G. and Bissel,M.J. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 1419-1423.
17. Bissell,M.J., Lee,E.Y.-H., Li,M.-L., Chen,L.-H. and Hall,H.G. (1985) The Second NIADDK Symposium on the Study of Benign Prostatic Hyperplasia, in press.
18. Sen,R. and Baltimore,D. (1986) Cell 46, 705-716.

19. Weinberger,J., Baltimore,D. and Sharp,P. (1986) Nature 322, 846-848.
20. Singh,H., Sen,R., Baltimore,D. and Sharp,P.A. (1986) Nature 319, 154-158.
21. Hennighausen,L., Siebenlist,U., Danner,D., Leder,P. Rawlins, D., Rosenfeld,P. and Kelly,T. (1985) Nature 314, 289-292.
22. Sive,H.L. and Roeder,R.G. (1986) Proc Natl. Acad. Sci. U.S.A. 83, 6382-6386.
23. Hatamochi,A., Paterson,B. and de Crombrugghe,B. (1986) J. Biol. Chem. 261, 11310-11314.
24. Heberlein,U., England,B. and Tjian,R. (1985) Cell 41, 965-977.
25. Dynan,W.S. and Tjian,R. (1985) Nature 316, 774-778.
26. Topol,J., Ruden,D.M. and Parker,C.S. (1985) Cell 42, 527-537.
27. Wu,C. (1985) Nature 317, 84-87.
28. Siebenlist,U., Hennighausen,L., Battey,J. and Leder,P. (1984) Cell 37, 381-391.
29. Borgmeyer,U., Nowock,J. and Sippel,A.E. (1984) Nucleic Acids Res. 12, 4295-4311.
30. Pastorcic,M., Wang,H., Elbrecht,A., Tsai,S.Y., Tsai,M.-J. and O'Malley,B.W. (1986) Mol. Cell. Biol. 6, 2784-2791.
31. Karin,M., Haslinger,A., Holtgreve,H., Richards,R.I., Krauter,P., Westphal,H.M. and Beato,M. (1984) Nature 308, 513-519.
32. Emerson,B.M., Lewis,C.D. and Felsenfeld,G. (1985) Cell 41, 21-30.
33. Sive,H., Heintz,N. and Roeder,R.G. (1986) Mol. Cell. Biol. 6, 3329-3340.
34. Prywes,R. and Roeder,R.G. (1986) Cell 47, 777-784.
35. Staudt,L.M., Singh,H., Sen,R., Wirth,T., Sharp,P. and Baltimore,D. (1986) Nature 323, 640-643.
36. Zinn,K. and Maniatis,T. (1986) Cell 45, 611-618.
37. Treisman,R. (1986) Cell 46, 567-574.
38. Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Nucleic Acids Res. 11, 1475-1489.
39. Fried,M. and Crothers,D. (1981) Nucleic Acids Res. 6, 6505-6525.
40. Garner,M. and Revzin,A. (1981) Nucleic Acids Res. 9, 3047-3060.
41. Vocke,C. and Bastia,D. (1983) Cell 35, 495-502.
42. Shalloway,D., Kleinberger,T. and Livingston,D.M. (1980) Cell 20, 411-422.
43. Bohnlein,E. and Gruss,P. (1986) Mol. Cell. Biol. 6 1401-1411.
44. Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambom,P. (1985) EMBO J. 4, 3563-3570.
45. Carthew,R.W., Chodosh,L.A. and Sharp,P.A. (1985) Cell 43, 439-448.
46. Miksicek,R., Heber,A., Schmid,W., Danesch,U., Posseckert,G., Beato,M. and Schutz,G. (1986) Cell 46, 283-290.