MOLECULAR AND CELLULAR BIOLOGY, Sept. 1990, p. 4690–4700
0270-7306/90/094690-11$02.00/0

Vol. 10, No. 9

Regulatory Elements Controlling Pituitary-Specific Expression of the Human Prolactin Gene

BERNARD PEERS, MARIANNE L. VOZ, PHILIPPE MONGET,† MARIANNE MATHY-HARTERT, MONIQUE BERWAER, ALEXANDRA BELAYEW, AND JOSEPH A. MARTIAL*

Laboratoire de Biologie Moléculaire et de Génie Génétique, Université de Liège, Institut de Chimie B6, B-4000 Sart-Tilman, Belgium

Received 16 April 1990/Accepted 25 June 1990

We have performed transfection and DNase I footprinting experiments to investigate pituitary-specific expression of the human prolactin (hPRL) gene. When fused to the chloramphenicol acetyltransferase (CAT) reporter gene, 5,000 base pairs of the 5′-flanking sequences of the hPRL gene were able to drive high *cat* gene expression in prolactin-expressing GH3B6 cells specifically. Deletion analysis indicated that this pituitary-specific expression was controlled by three main positive regulatory regions. The first was located just upstream from the TATA box between coordinates −40 and −250 (proximal region). We have previously shown that three motifs of this region bind the pituitary-specific Pit-1 factor. The second positive region was located in the vicinity of coordinates −1300 to −1750 (distal region). DNase I footprinting assays revealed that eight DNA motifs of this distal region bound protein Pit-1 and that two other motifs were recognized by ubiquitous factors, one of which seems to belong to the AP-1 (*jun*) family. The third positive region was located further upstream, between −3500 and −5000 (superdistal region). This region appears to enhance transcription only in the presence of the distal region.

Many genes of pluricellular organisms are expressed in only one differentiated cell type. Control of this tissue-specific expression is mostly transcriptional. In recent years, many transfection studies with chimeric genes have shown that regulatory sequences controlling gene transcription are often located just upstream from the transcription initiation (CAP) site (11), but can also be found very far upstream or downstream from the CAP site (2, 28, 41). Sensitive DNA-binding assays, such as DNase I footprinting (13) and gel retardation (12), have revealed that these *cis*-acting elements contain DNA sequence motifs which bind specific *trans*-acting proteins (for a review, see reference 18). Interestingly, while some of these regulatory proteins are present in many cell types and stimulate transcription of several genes (e.g., Sp1, AP-1, Oct-1) (1, 5, 11, 32), others are found only in one particular cell type and are involved in the tissue-specific expression of the gene to which they bind (e.g., Oct-2) (34, 39).

In mammals, prolactin (PRL) is expressed primarily by the lactotrophic cells of the anterior pituitary. Several in vivo and in vitro studies performed on the rat PRL gene have indicated that 400 base pairs (bp) of its 5′-flanking sequence are sufficient to confer pituitary-specific expression (6, 16, 36, 40). This is achieved through the binding to this region of a pituitary-specific factor named LSF-1 (16, 20), Puf-1 (6, 7, 42), or Pit-1 (27, 36). For simplicity's sake, this factor will be referred to as Pit-1. A pituitary-specific enhancer was also identified between coordinates −1530 and −1800 in the rat PRL gene by Nelson et al. (36, 37). This distal DNA region was shown to contain four Pit-1-binding sites (27, 36). However, while this enhancer was described by these authors as the major determinant of rat PRL expression, other studies have failed to identify this enhancer (6, 26).

Recently, the cDNA coding for the Pit-1 protein has been cloned by Ingraham et al. (17). They also demonstrated that Pit-1 was able to bind and to activate the promoters of both the rat prolactin and growth hormone genes (17, 27). Independently, Bodner et al. cloned exactly the same cDNA (4). In their experiments, however, the encoded protein (named GHF-1) appeared to bind and activate the growth hormone promoter but not the PRL promoter (4, 10, 43). What causes this discrepancy is unclear.

We have previously reported the cloning and sequencing of the human PRL gene (44). Recently, using a cell-free in vitro transcription system, we have shown that 740 bp of 5′-flanking sequences of the human PRL gene drive efficient transcription only in extracts of pituitary-derived GH3B6 cells. By DNase I footprinting experiments, we also identified three Pit-1-binding sites in the proximal 250 bp of the human PRL promoter (24).

In this article, we show by transfection experiments that additional regulatory DNA regions, located far upstream in the human PRL gene, also play an important role in its pituitary-specific transcription. By DNase I footprinting assays, we demonstrate that one of these upstream regions contains numerous binding sites for pituitary-specific and ubiquitous factors. Finally, competition experiments suggest that all the pituitary-specific footprints are due to the binding of protein Pit-1 and that one of the ubiquitous factors belongs to the AP-1 protein family.

MATERIALS AND METHODS

Plasmid constructions. Plasmids RSVCAT (14), CMV CAT, p740CAT (24), and TKCAT (25) have been described previously. The RSVamp plasmid contains the Rous sarcoma virus enhancer/promoter inserted upstream from the β-lactamase RTEM (Amp^r) gene, followed by the simian virus 40 splicing and polyadenylation signals (33).

The p5000CAT plasmid, containing 5 kb of the human PRL gene 5′ region and 14 base pairs of the transcribed sequence fused to the bacterial chloramphenicol acetyltrans-

* Corresponding author.
† Present address: INRA, Station de Physiologie de la Reproduction, URA CNRS 234, Nouzilly, 37380 Monnaie, France.

ferase (CAT) gene, was constructed as follows. A 4.25-kb *Hin*dIII DNA fragment from the genomic recombinant phage EMBLhPRL3 (Berwaer et al., unpublished), containing 5'-flanking sequences of the human PRL gene (from coordinates −740 to −5000), was gel purified and cloned into the *Hin*dIII site of the p740CAT plasmid. The inserted fragment was obtained in both orientations, yielding plasmid p5000 CAT for the correct orientation and p(740-5000)740CAT for the inverted orientation (see Fig. 2). The internal and 5' deletion mutants presented in Fig. 2A were prepared from plasmid p5000CAT by appropriate restriction digestion followed by blunt-ending with the Klenow fragment of DNA polymerase I and by recircularization with T4 DNA ligase. The structures of all mutants were confirmed by restriction mapping.

All plasmids were prepared by alkaline lysis, purified by centrifugation in CsCl-ethidium bromide gradients, treated with RNase, centrifuged through a cushion of 1 M NaCl, extracted with phenol-chloroform, and finally precipitated with ethanol. The concentrations of all plasmid preparations were determined by measuring the absorbance at 260 nm.

Cell culture and electroporation. GH3B6 cells deriving from a rat pituitary tumor (22) and HeLa cells were grown in monolayer in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal calf serum (FCS). For transfection assays, cells were harvested with trypsin-EDTA and resuspended in DMEM medium with 10% FCS at a concentration of 20×10^6 cells per ml. Purified plasmid DNA (5 pmol for the PRL mutant constructs and TKCAT, 1 pmol for RSVCAT and CMVCAT) were mixed with 16×10^6 cells in the presence of 1 pmol of plasmid RSVamp (used as a control for DNA transfection efficiency). The cells were then exposed to a single pulse of 250 V/4 mm and 1,500 μF capacitance for GH3B6 cells or 300 V/4 mm and 1,050 μF for HeLa cells, with a Celljject apparatus (SEDD, Liège, Belgium). The transfected cells were immediately transferred into four tissue culture dishes (35 mm) and maintained in DMEM with 10% FCS. After 48 h, the cells were harvested by scraping, washed in phosphate-buffered saline, suspended in 100 μl of 50 mM sodium phosphate buffer (pH 6.0), and frozen at −70°C. The cells were disrupted by three cycles of freezing/thawing with sonication, and the extracts were centrifuged at 12,000 × *g* for 15 min in an Eppendorf centrifuge to remove cell debris. Then, 40 μl of supernatant was used for the β-lactamase assay and 50 μl for the CAT assay.

β-Lactamase assay. The β-lactamase assay is based on the change of the 482-nm absorption spectrum of Nitrocefin upon hydrolysis by β-lactamase (38). Cell extract supernatant (40 μl) was added to 400 μl of 100 μM Nitrocefin (Oxoid, Basingstoke, England) and to 60 μl of 50 mM sodium phosphate buffer (pH 6). The reaction mixtures were incubated at 4°C in the dark. The reaction was stopped by addition of 50 μl of 10% sodium dodecyl sulfate, and β-lactamase activity was determined by measuring the absorbance at 482 nm.

CAT assay. Cell extract supernatant (50 μl) was incubated for 6 min at 65°C and centrifuged for 15 min at 12,000 × *g*. Then, 5 μl of [^{14}C]chloramphenicol (25 μCi/ml; Amersham), 11 μl of acetyl coenzyme A (40 mM; Sigma), 40 μl of Tris hydrochloride (Tris-HCl; 1 M, pH 7.8), and 82 μl of H_2O were added to the supernatants. The reactions were stopped by addition of ethyl acetate. Chloramphenicol and its acetylated forms were separated by thin-layer chromatography (15). Following autoradiography, the silica gel spots from the thin-layer chromatography plates containing either acetylated products or remaining substrate were scraped into separate scintillation vials and counted. CAT activity was normalized for amp expression. Data are expressed relative to RSVCAT activity, which was set at 100% (see Fig. 2).

Extracts and DNase I footprinting. Whole-cell extracts (WCE) were prepared as described by Manley (29). Protein concentrations of these extracts, usually 10 to 20 mg/ml, were determined by the Bradford assay with bovine serum albumin as a standard.

DNase I footprinting reactions were performed as described by Jones et al. (19). End-labeled DNA probes (1 to 2 ng), prepared by treatment with T4 polynucleotide kinase, were incubated with the cell extract (up to 150 μg of protein) in a final volume of 50 μl containing 10% (vol/vol) glycerol, 1% polyvinylalcohol, 25 mM Tris-HCl (pH 7.9), 6.25 mM $MgCl_2$, 50 mM KCl, 1 mM EDTA, 0.5 mM dithiothreitol, and 1 μg of poly(dI-dC). After 15 min at 0°C followed by 2 min at 20°C, 50 μl of a 5 mM $CaCl_2$ solution plus 10 mM $MgCl_2$ was added. Digestion with 1.25 to 25 ng of DNase I (Boehringer) was then allowed to proceed at 20°C for 1 min. The reaction was stopped and analyzed on 5% polyacrylamide–42% urea gels, as described by Jones et al. (19).

Oligonucleotides and competition experiments. The oligodeoxyribonucleotides were obtained from Eurogentec (Liège, Belgium). Oligonucleotide GH1 contains the proximal GHF-1 (Pit-1)-binding site of the human growth hormone promoter (23, 24). Oligonucleotides P1 and P2 respectively contain the first and second Pit-1-binding site of the human PRL promoter (24). The oligonucleotide AP-1 contains the AP-1 site of the human collagenase promoter (1). Oligonucleotide C/EBP was described by Landschulz et al. (21). The sequences of these oligonucleotides are shown below (only one strand is presented).

GH1	GATCCCATGCATAAATGTACACAG
P1	CTAGAATGCCTGAATCATTATATTCATGAAGATATC
P2	GATCCCTTCCTGAATATGAATAAGAAATAAAAG
P2mut	GATCCCTTCCTGAATACAAATAAGAAATAAAAG
AP-1	AGCTTAAAGCATGAGTCAGACACCT
C/EBP	AATTCAATTGGGCAATCAGG
D8	GATCCAAGAATGTTGCAACACCTTTA

For the competition experiments, the labeled DNA fragments were incubated with the protein extract in the presence of various amounts of double-stranded oligonucleotides. DNase I digestion was then performed as in the DNase I footprinting experiments.

RESULTS

Comparison of the rat and human PRL 5'-flanking sequences. We have previously reported the isolation of the human PRL gene containing 900 bp of proximal 5'-flanking sequences (44). Recently, additional upstream regions have been cloned and the nucleotide sequence was determined up to coordinate −3470 (Berwaer et al., unpublished).

The 5'-flanking sequences of the human and rat PRL genes were then compared. Figure 1 shows the dot plot matrix comparison, which revealed the presence of five regions of high similarity between the two upstream sequences (boxes I to V, Fig. 1). Box I, representing 71% similarity, was located in the proximal 470 bp of the PRL promoters. We and others have previously shown that this region of the human and rat gene is important for promoter activity and binds the *trans*-acting factor Pit-1 (6, 16, 24, 36, 40). Box III, representing 65% similarity, was located between coordinates −1500 and −1800 in the rat PRL gene. Nelson et al. reported the presence, in this region, of a pituitary-specific

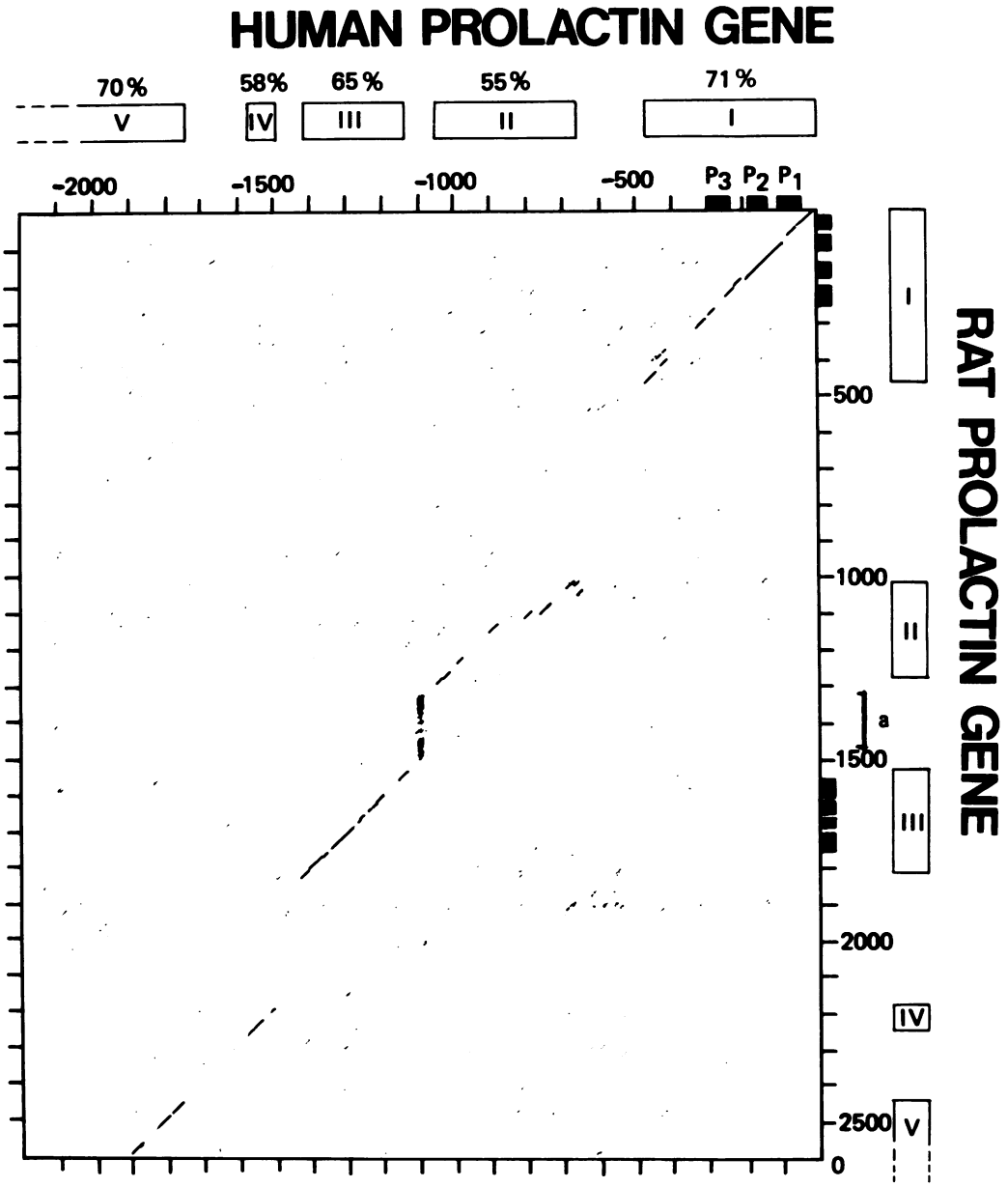


FIG. 1. Dot plot matrix comparison of rat and human PRL gene 5'-flanking sequences. A 2,600-bp portion of rat PRL 5' sequences (along the vertical axis) (31) and of human PRL 5' sequences (along the horizontal axis) (Berwaer et al., unpublished) were compared by using the program Compare from the University of Wisconsin Genetics Computer Group (9). Numbers along the two axes refer to the coordinates from the CAP site. The Pit-1-binding sites in the rat PRL 5' sequences (36) and in the human PRL promoter (P1, P2, and P3) (24) are represented by black boxes. Numbered boxes (I to V) indicate the regions of high sequence similarity. The vertical line drawn in the center of the matrix is due to repetitive TGTG sequence motifs present in the rat 5' sequences between coordinates −1300 and −1450 (bracket a).

enhancer containing four Pit-1-binding sites (8, 26, 36, 37). Note that box III was closer to the CAP site in the human PRL gene, i.e., between coordinates −1100 and −1400. Other homologous regions were present further upstream (box IV, 58% similarity; box V, 71% similarity). The comparative analysis could not extend beyond −2600 for lack of further rat sequence data.

This sequence comparison shows that several 5' regions of the PRL gene have been conserved during evolution, suggesting that they could be functionally important and could contain *cis*-acting regulatory elements.

Pituitary-specific transcription of the human PRL gene is controlled by proximal and distal 5' regions. To investigate the presence of regulatory regions in the 5'-flanking sequences of the human PRL gene, we fused 5,000 bp of the human PRL 5' sequences to the *cat* reporter gene (plasmid p5000CAT, Fig. 2A). Several 5' and internal deletions were then generated in the human PRL 5' sequences, as shown in Fig. 2A. All of these constructs were introduced by electroporation into PRL-expressing GH3B6 cells and PRL-nonexpressing HeLa cells. In all the transfection experiments, cells were cotransfected with an internal control plasmid, RSVamp (33), in order to eliminate variation of CAT activity due to possible differences in transfection efficiency.

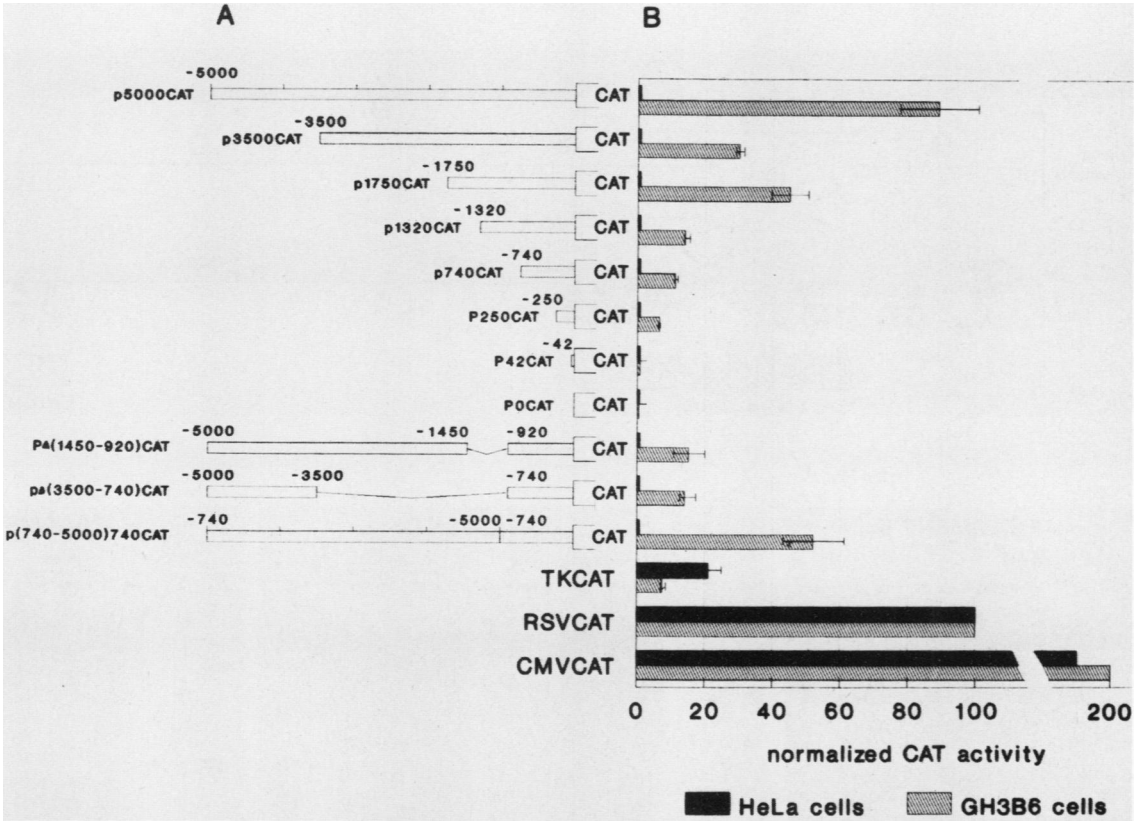


FIG. 2. Analysis of transient CAT expression obtained with the human PRL-*cat* fusion genes in GH3B6 and HeLa cells. (A) Schematic representation of the 5′ and internal human PRL mutant promoters fused to the *cat* reporter gene. (B) CAT expression in GH3B6 and HeLa cells. GH3B6 and HeLa cells were cotransfected by electroporation with one of the constructs presented in panel A and with 1 pmol of plasmid RSVamp (see Materials and Methods). After 48 h, CAT and β-lactamase enzymatic activities were determined in the transfected cell extracts. Panel B presents CAT activity normalized to β-lactamase activity. The values are expressed as a percentage of the CAT/β-lactamase activity ratio obtained with the RSVCAT plasmid, which was arbitrarily set to 100%. The results shown are averages of five separate transfection experiments for GH3B6 cells and two separate experiments for HeLa cells and have been reproduced with several plasmid preparations. As the transfection efficiency was very similar for all the constructs, the results expressed with or without β-lactamase normalization were virtually identical (data not shown).

Figure 2B presents the normalized CAT activity obtained with each construct after transfection of GH3B6 and HeLa cells. Several observations can be made.

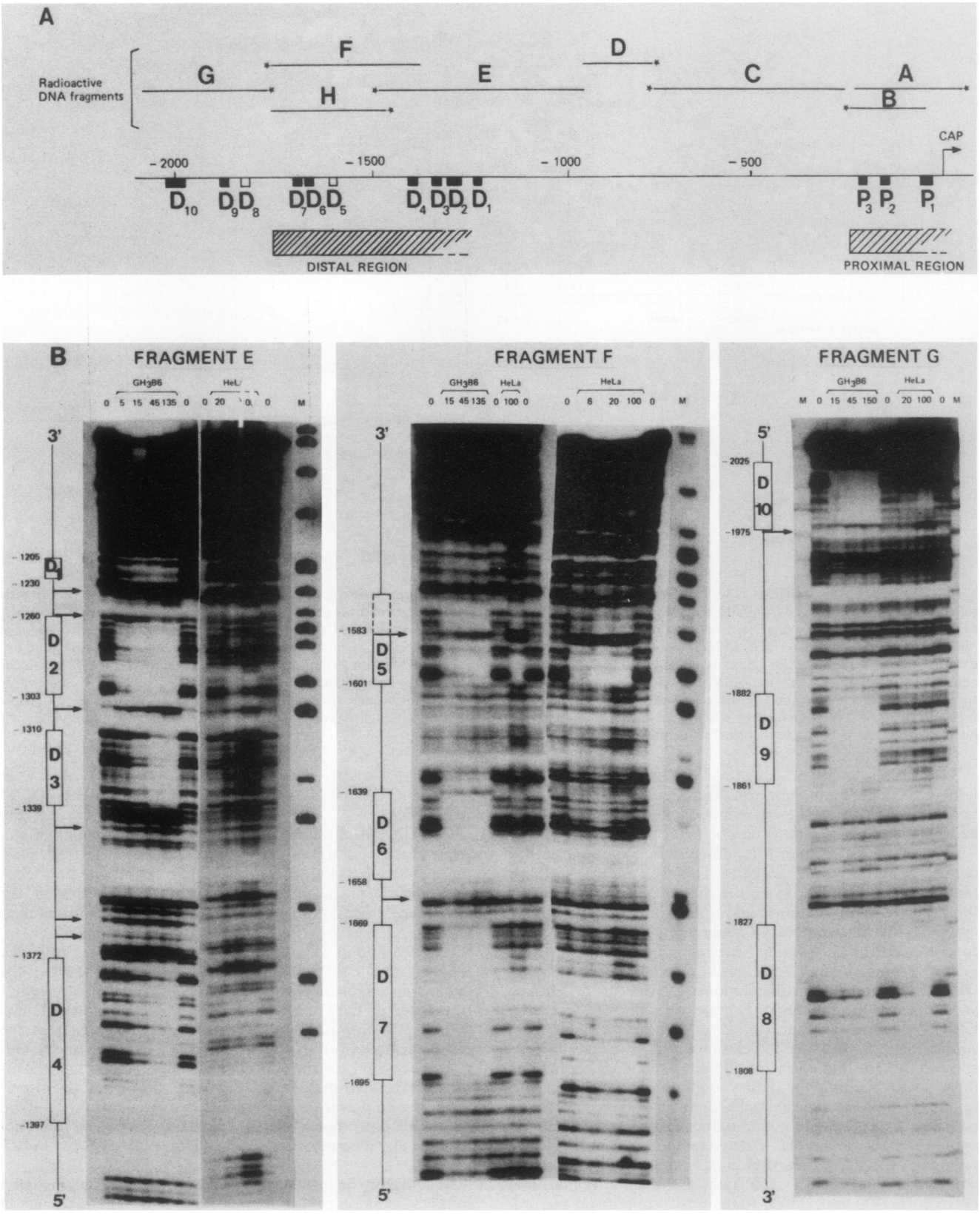
(i) None of the human PRL constructs generated detectable CAT activity in HeLa cells. This was not due to a low transfection efficiency, since the three control promoters—the herpesvirus thymidine kinase, Rous sarcoma virus, and cytomegalovirus promoters—generated strong CAT signals. In contrast, the 5′-flanking sequences of the human PRL gene drove significant CAT expression when introduced into GH3B6 cells. These results demonstrate that the 5′-flanking sequences of the human PRL gene are sufficient to confer tissue-specific expression.

(ii) Transfection of the 5′ deletion mutants in GH3B6 cells showed that the human PRL 5′-flanking sequences contained three main positive regulatory regions. The first positive region, located in the proximal promoter sequences between coordinates −42 and −250, enhanced promoter activity approximately sevenfold. The second and third positive regions were located farther upstream, in the vicinity of coordinates −1320 to −1750 (named the distal region) and between −3500 and −5000 (named the superdistal region). Each of these two upstream regions enhanced transcription approximately three- to fourfold. The region from −1750 to −3500 had a slight but reproducible inhibitory effect on CAT expression.

(iii) The distal and superdistal regulatory regions stimulated transcription in the inverted orientation. Indeed, plasmid p(740-5000)740CAT was 4.5-fold more active than plasmid p740CAT, but stimulation was slightly less in the antisense orientation than in the natural orientation (p5000CAT was eightfold more active than p740CAT).

(iv) Deletion of sequences −900 to −1450, which removed a part of the distal regulatory region, virtually abolished the enhancing effect of both the distal and superdistal regions [the level of CAT expression was the same for pΔ(1450-900)CAT as for p740CAT]. This result indicates that the distal region has an essential regulatory function and seems to be necessary for the action of the superdistal region. This hypothesis is strengthened by the fact that when the superdistal region was inserted directly upstream from the 740-bp human PRL promoter [plasmid pΔ(3500-740)CAT], it did not by itself significantly enhance promoter activity (see Discussion).

***trans*-Acting factors bind to sequences located far upstream in the 5′-flanking part of human PRL gene.** The presence of regulatory regions in the 5′-flanking part of the human PRL gene suggests that these 5′ sequences could be recognized by


A
Radioactive DNA fragments
G
F
H
E
D
C
A
B
CAP
-2000
-1500
-1000
-500
D10 D9 D8 D7 D6 D5 D4 D3 D2 D1
P3 P2 P1
DISTAL REGION
PROXIMAL REGION
B
FRAGMENT E
FRAGMENT F
FRAGMENT G
GH386
HeLa
3'
5'
D1
D2
D3
D4
D5
D6
D7
D8
D9
D10

trans-acting factors. Therefore, a systematic DNase I footprinting study was performed on the 5'-flanking part of the human PRL gene with GH3B6 and HeLa WCE. Figure 3A depicts the different ^{32}P-labeled DNA fragments (A to H) encompassing the first 2,060 bp of human PRL 5' sequences used in this study. The three footprints observed in the proximal promoter sequences (P1, P2, and P3) due to the binding of protein Pit-1 have been described previously (24). No footprints were observed with fragments D and C in the sequences between −250 and −920 (data not shown). However, with fragments E to H, located more upstream, several DNase I-protected sequences were detected.

With fragment E, four footprints, situated between coordinates −1200 and −1400, were observed in the presence of GH3B6 cell extracts (Fig. 3B, D1 to D4). DNase I-hypersensitive sites also appeared between these sites (arrows in Fig. 3B and 4A). These footprints were not observed in the presence of HeLa WCE, indicating that the *trans*-acting factors binding to these four sequences were only present (or active) in GH3B6 cells. Site D2 was protected at a lower protein extract concentration than the other sites, suggesting either that the corresponding *trans*-acting factor is more abundant or that it binds to the D2 sequences with a higher affinity. Note that these four distal GH3B6-specific footprints were located in homologous region III described in Fig. 1.

When DNase I footprinting assays were performed on fragment F (−1350 to −1750), three additional footprints were observed (Fig. 3B, D5 to D7). Footprints D6 and D7 appeared only with GH3B6 extracts, while D5, characterized by a very strong hypersensitive site at its 3' border, was observed with both GH3B6 and HeLa WCE. Sequences just downstream from site D5 were also slightly protected, suggesting lower-affinity binding of a protein(s) (hatched box in Fig. 3B).

Finally, fragment G (−1750 to −2060) yielded three additional footprints (Fig. 3B, D8, D9, and D10). Footprint D8 was obtained with HeLa extracts and with high GH3B6 WCE concentrations. Extracts of rat liver or BeWo cells (human placental cells) also produced this footprint (data not shown). This indicates that the *trans*-acting factor binding to site D8 is ubiquitous but is less abundant (or less active) in GH3B6 cells. Footprints D9 and D10 were observed only with GH3B6 WCE. Nevertheless, high HeLa WCE concentrations did slightly protect D10 sequences, but the footprint pattern differed from that observed with GH3B6 WCE. This suggests that high-affinity binding occurs with a *trans*-acting factor(s) present only in GH3B6 WCE, but that other proteins present in HeLa WCE (and other heterologous WCE; data not shown) may bind with a lower affinity (see Discussion).

All these footprinting results have been reproduced with several different cell extract preparations. They clearly demonstrate that GH3B6-specific and ubiquitous *trans*-acting factors bind near and within the distal regulatory region of the human PRL gene. Figure 3A summarizes the locations of all the footprints observed within the gene.

Identification of the *trans*-acting factors binding to the upstream human PRL sequences. (i) Sequence analysis of the DNase I-protected sites. To identify the *trans*-acting factors responsible for all the footprints described above, we first analyzed the sequences of the protected sites. As shown in Table 1, the TATNCAT motif, where N is any nucleotide, previously proposed to be the Pit-1 consensus binding sequence (24, 36), was found in all GH3B6-specific footprints, either perfectly conserved (D2 and D6) or with one (D1, D3, D4, and D10) or two (D7 and D9) mismatches. The larger D2 and D10 sites contained two and three related consensus sequences, respectively. Furthermore, most of these GH3B6-specific footprints contained A+T-rich sequences, another characteristic of Pit-1-binding sites. This suggests that all the GH3B6-specific footprints may be due to the binding of protein Pit-1.

The sequence of the ubiquitous D5 footprint contained the ATGAATCAT motif, which is similar to the AP-1 consensus binding site [ATGA(G\C)TCAT] (5). Finally, the sequence of the ubiquitous D8 footprint revealed a perfect 10-bp palindrome (underlined in Table 1). To our knowledge, no such palindromic sequence has been reported so far as a *trans*-acting factor-binding site.

(ii) Oligonucleotide competition experiments. To investigate further the identity of the proteins responsible for the different footprints observed, DNase I footprinting experiments were performed in the presence of competing amounts of various synthetic double-stranded oligonucleotides (see Materials and Methods for the sequences of these oligonucleotides). As shown in Fig. 4A, the four distal GH3B6-specific footprints observed with fragment E (D1 to D4) disappeared in the presence of 50 ng of oligonucleotides P1 and GH1, which contain the respective proximal Pit-1-binding sites of the human PRL and growth hormone promoters (24) (Fig. 4A, lanes 3 and 5). The two other competing oligonucleotides, containing the D8 10-bp palindromic sequence or containing a high-affinity C/EBP-binding site, had no influence on these four footprints (lanes 4 and 6). These results strongly support the view that Pit-1 is the factor that binds to the four distal sites.

Similarly, the D6 and D7 sequences were no longer protected by GH3B6 WCE in the presence of 50 ng of oligonucleotide GH1 (Fig. 4B, lane 4) or oligonucleotide P1 (data not shown). The two other oligonucleotides, containing a high-affinity AP-1-binding site or the D8 palindromic sequence, were unable to compete for binding to sites D6 and D7 (Fig. 4B, lanes 3 and 5). Inversely, D5 site protection was not prevented by oligonucleotide GH1 or D8, but disappeared upon addition of 50 ng of oligonucleotide AP-1. Note that the D5-associated DNase I-hypersensitive site also disappeared with oligonucleotide AP-1 but not with the two

FIG. 3. DNase I footprinting analysis performed on the first 2,060 bp of the human PRL gene 5'-flanking sequences. (A) Schematic representation of the different ^{32}P-labeled DNA fragments encompassing the 2,060 bp of the human PRL 5'-flanking sequences used in the footprinting experiments (fragments A to H). Solid and open boxes indicate the respective locations of pituitary-specific and ubiquitous footprints observed in this study. The two hatched boxes show the approximate positions of the proximal and distal regulatory regions identified by the transfection experiments (see Fig. 2). (B) DNase I footprinting analysis of the labeled DNA fragments E, F, and G with GH3B6 and HeLa WCE. Numbers above each lane indicate the amount of protein extract added (in micrograms). DNA fragment labeling and DNase I protection reactions were carried out as described in Materials and Methods. Open boxes alongside the autoradiographs indicate the DNase I-protected sites (D for distal) with the coordinates of their respective 5' and 3' borders. Arrows show the DNase I-hypersensitive sites. Fragment E (−1450 to −920) was end labeled at the *Nde*I site (−1450). Fragment F (−1750 to −1350) and fragment G (−2060 to −1750) were labeled at the *Afl*II site (−1750). The markers (lane M) are pBR322 digested by *Hin*fI, except for fragment E, where the marker is pBR325 digested by *Hpa*II.

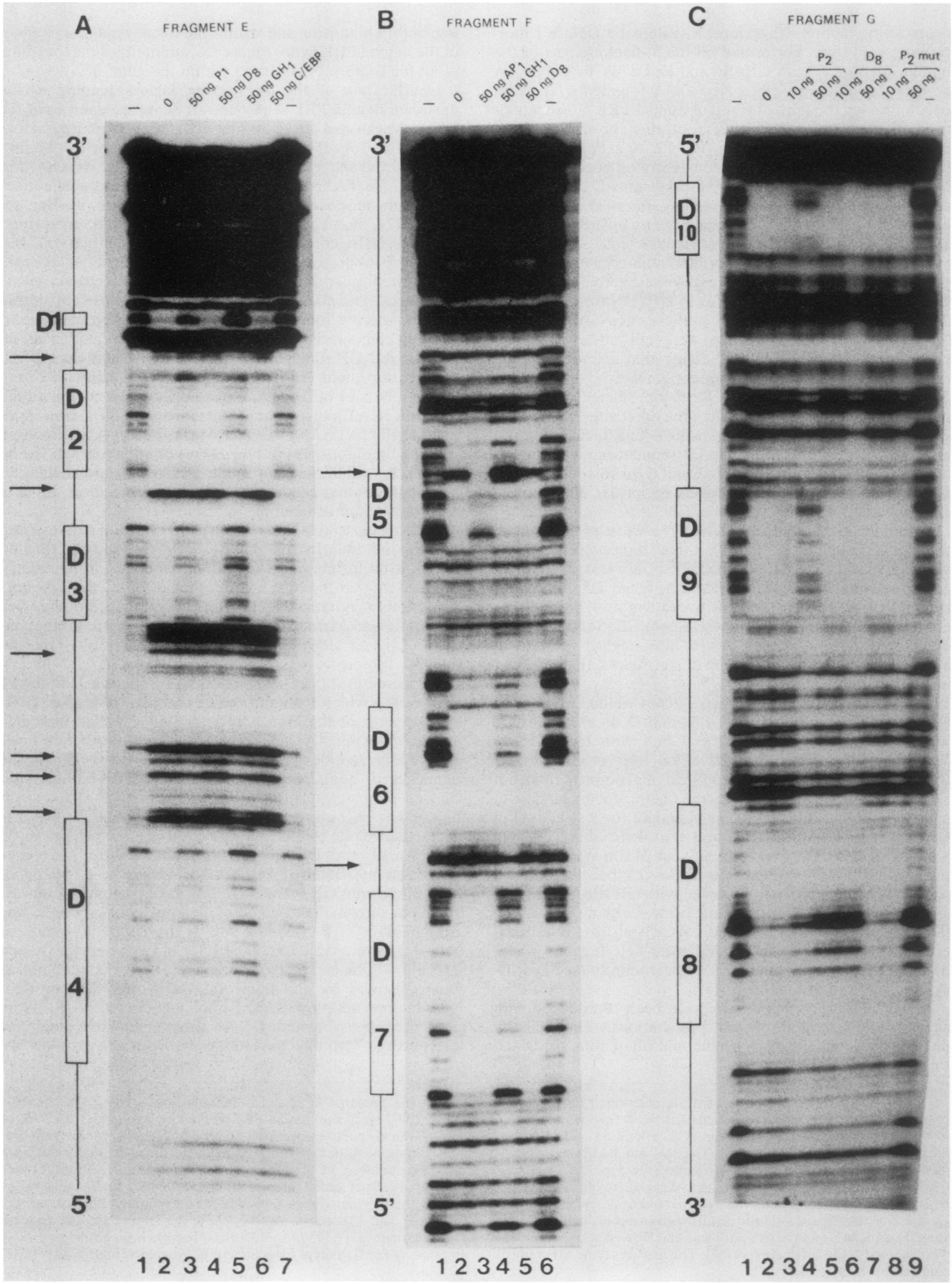

A
FRAGMENT E
— 0 50 ng P1 50 ng D8 50 ng GH1 50 ng C/EBP —
3'
D1
D 2
D 3
D 4
5'
1 2 3 4 5 6 7
B
FRAGMENT F
— 0 50 ng AP1 50 ng GH1 50 ng D8 —
3'
D 5
D 6
D 7
5'
1 2 3 4 5 6
C
FRAGMENT G
P2 D8 P2 mut
— 0 10 ng 50 ng 10 ng 50 ng 10 ng 50 ng —
5'
D 10
D 9
D 8
3'
1 2 3 4 5 6 7 8 9

TABLE 1. Sequences of DNase I-protected sites in human PRL gene 5′ region[a]

A.	Sequences of the eight GH3B6-specific footprints.		
D1	(-1230)	ATTTTTAAGCCCATGAAAAGATGAATTT TAAAAATTCGGGTACTTTTCTACTTAAA	(-1205)
D2	(-1303)	AAGTCATCAATAAAACTTAGTTCTTATTCATATAGTCATTTCA TTCAGTAGTTATTTTGAATCAAGAATAAGTATATCAGTAAAGT	(-1260)
D3	(-1339)	GGCTGCTTTAGATGCATGTACTTAAAAAT CCGACGAAATCTACGTACATGAATTTTTA	(-1310)
D4	(-1397)	TAAATTACCTCATTCATCAGTCTATG ATTTAATGGAGTAAGTAGTCAGATAC	(-1372)
D6	(-1658)	AAAAATGAATACCTAATGTG TTTTTACTTATGGATTACAC	(-1639)
D7	(-1695)	TGCTTTCTTAATAATTCAGACTCTCTC ACGAAAGAATTATTAAGTCTGAGAGAG	(-1669)
D9	(-1882)	CAGACCCTGATAAACACAGCTT GTCTGGGACTATTTGTGTCGAA	(-1861)
D10	(-2026)	TTTTTGCATAATATATGTCTTTGCATTATTTATATATTTCAATATTCCATTC AAAAACGTATTATATACAGAAACGTAATAAATATATAAAGTTATAAGGTAAG	(-1975)
Pit-1 consensus : 5′-TATNCAT-3′			
B.	Sequences of the two ubiquitous footprints		
D5	(-1601)	TCTATGAATCATCACAT AGATACTTAGTAGTGTA	(-1585)
AP1 consensus :5′-ATGA(G/C)TCAT-3′			
D8	(-1828)	CCAAGAATGTTGCAACACCTT GGTTCTTACAACGTTGTGGAA	(-1808)

[a] Pit-1 (36) and AP-1 (5) consensus sequences are indicated by solid and open arrows, respectively. The 10-bp palindromic sequence of the D8 site is underlined.

other oligonucleotides. Likewise, footprint D5 obtained with HeLa WCE disappeared only with oligonucleotide AP-1 (Fig. 5A, lane 8). These competitions strongly suggest that protein Pit-1 binds to the D6 and D7 sites and that site D5 is recognized in GH3B6 and HeLa WCE by a protein belonging to the AP-1 family.

Figure 4C depicts analogous competition experiments performed on fragment G. Oligonucleotide P2 (50 ng), which contains the second Pit-1-binding site of the human PRL promoter, competed for binding to sites D9 and D10 (lane 4). These two GH3B6-specific footprints also disappeared completely with 50 ng of oligonucleotides GH1 and P1 (data not shown).

To ascertain further that sequences D9 and D10 bound the Pit-1 factor and not another A+T-rich sequence-binding factor, we synthesized a mutated P2 oligonucleotide in which two base substitutions were introduced within the P2 site sequence. This mutated P2 oligonucleotide (P2mut), which no longer bound Pit-1 but which retained most of the A+T-rich sequence of site P2, failed to compete for binding to sequences D9 and D10 (lanes 7 and 8).

As outlined above, the D8 sequences were not completely protected by GH3B6 WCE (Fig. 4C, lane 2). However, we observed that D8 site protection disappeared only with the 10-bp palindrome-containing D8 oligonucleotide; this competition was more obvious with HeLa WCE, as shown in Fig. 5B (lane 5).

DISCUSSION

Three regulatory regions control pituitary-specific expression of the human PRL gene. In this study, we have performed transfection and DNase I footprinting experiments to investigate pituitary-specific expression of the human PRL gene. A 5,000-bp portion of 5′-flanking sequences of the human PRL gene was able to drive high expression of the *cat* reporter gene in PRL-expressing GH3B6 cells but not in PRL-nonexpressing HeLa cells. Deletion studies have revealed that three main positive regulatory regions are involved in this pituitary-specific expression.

The first region is located in the proximal promoter sequences between coordinates −250 and −40. This is in agreement with our previous report that the 740-bp human PRL promoter exhibits a pituitary-specific activity in a cell-free transcription system (24). This activity correlates with the presence in this proximal region of three binding sites for the pituitary-specific factor Pit-1 (Fig. 3A, P1, P2, and P3).

The sequence between coordinates −250 and −1200 had little effect on transcription (Fig. 2), and no DNase I footprint was detected in this region with either GH3B6 or HeLa WCE. The second positive regulatory region lay further upstream, between coordinates −1750 and −1320 (or possibly spanning −1320). This distal regulatory region contained many binding sites for both pituitary-specific and ubiquitous factors (Fig. 3A).

The third positive region was still further upstream, between coordinates −5000 and −3500. The presence of protein-binding sites in this superdistal region has not been investigated in this study. The ability of the distal and superdistal regions to stimulate transcription while located so far upstream from the CAP site and also in the inverse orientation (Fig. 2) is indicative of an enhancerlike function.

FIG. 4. Competition footprinting experiments performed on DNA fragments E, F, and G with GH3B6 WCE. The labeled DNA fragments were incubated with GH3B6 extract (90 μg of protein) in the presence of competing double-stranded oligonucleotides (see Materials and Methods). The amount (in nanograms) of competing oligonucleotide is noted above each lane. Control DNase I digestion without WCE is indicated (—). Arrows indicate DNase I-hypersensitive sites.

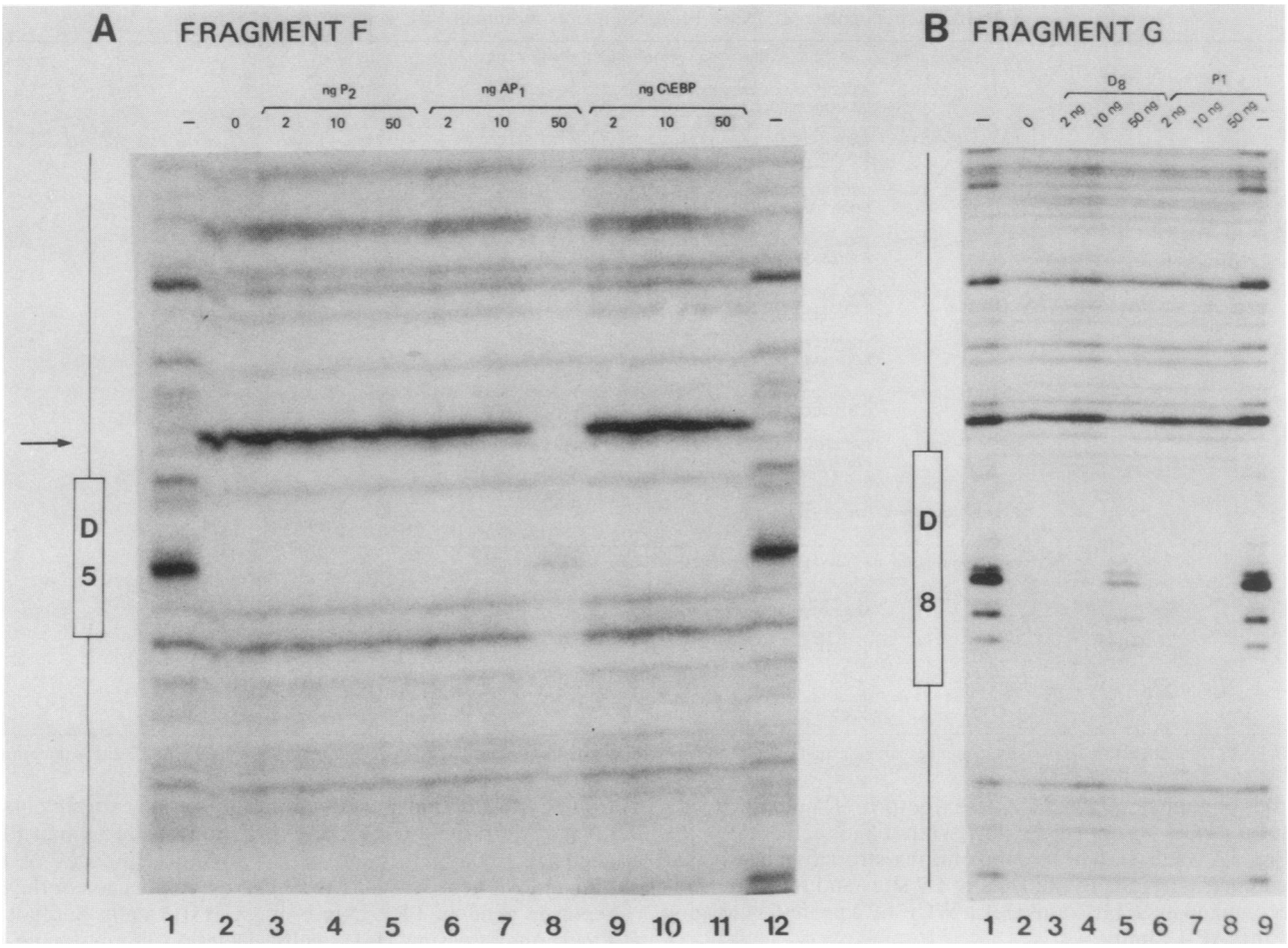


FIG. 5. Competition footprinting experiments performed on fragments F and G with HeLa WCE. The labeled DNA fragments were incubated with HeLa extract (100 μg of protein) in the presence of competing oligonucleotides (see Materials and Methods). The amount of competing oligonucleotide is indicated above each lane. Arrow indicates DNase I-hypersensitive site.

Since deletion of the distal region abolishes the enhancing effect of the superdistal region, it seems that the distal region is required for the action of the superdistal region, suggesting a possible interaction between these two regions.

Identification of the *trans*-acting factors binding to the human PRL gene. In this study, we show that pituitary-specific and ubiquitous factors bind near and within the distal regulatory region defined by deletion studies. Competition experiments suggest that all GH3B6-specific footprints are due to binding of the pituitary-specific factor Pit-1. Furthermore, since oligonucleotides P1 and P2, deriving from the human PRL promoter, and oligonucleotide GH1, deriving from the human growth hormone promoter, were equally effective competitors for binding to these sites (Fig. 4), our results agree with Rosenfeld's conclusion that factor Pit-1 binds both promoters (24, 27, 36). Another explanation could be that the pituitary-specific elements of the growth hormone and PRL genes are recognized by a family of Pit-1/GHF1-related proteins, as has been shown for the Oct protein family (18).

The TATNCAT motif has previously been proposed as the Pit-1 consensus binding site (36). A recent base substitution mutational study has strengthened this view (42). Among the eight Pit-1 binding sites observed in the human PRL distal region, only sites D2 and D6 contain the exact consensus sequence, while the others contain related sequences with one or even two mismatches (Table 1). Why does Pit-1 bind to these degenerate motifs? One hypothesis is that binding depends not only on the central consensus motif but also on the flanking residues, as recently reported for binding of factor OBP100. In this latter case, Baumruker et al. have shown that when the canonical octamer consensus is degenerate, OBP100 binding depends on the flanking residues (3). A second explanation may be that there is a cooperativity in the binding of several Pit-1 proteins to adjacent sites. DNase I-hypersensitive sites, often observed between Pit-1-binding sites (Fig. 3 and 4), could be due to this cooperativity, since interaction between proteins bound to their respective sites creates bends in the DNA, where DNase can easily generate nicks.

The high-affinity Pit-1-binding sites D2 and D10 are larger than the others and contain two and three related Pit-1 consensus motifs, respectively. This strongly suggests that two or three Pit-1 proteins bind to these sites. This view is further supported by footprinting results showing that the two sites are not uniformly protected (data not shown). In addition to Pit-1 binding in GH3B6 WCE, site D10 was also slightly protected by heterologous cell extracts (HeLa, BeWo, rat liver extracts; Fig. 3 and data not shown). This indicates that ubiquitous factors also bind to this site but with a much lower affinity. In fact, the D10 sequence contains the motif CTTTGCAT, which resembles the con-

sensus binding site of the ubiquitous factor Oct-1 (ATTTG CAT). As previously mentioned by others, the Pit-1 and Oct-1 consensus binding sequences are quite similar, and these two regulatory proteins, which contains both homeo- and POU domains (17, 27, 43), are able to cross-bind the two consensus motifs, albeit with different affinities (17). Oct-1 binding might thus explain the observed partial protection of the D10 site by heterologous cell extracts. There are many recently reported examples of one DNA sequence being able to bind multiple regulatory proteins (18).

We have also shown that two ubiquitous factors bind to the 5′ region of the human PRL gene (Fig. 3, sites D5 and D8). The D5 sequence analysis and competition experiments strongly suggest that the ubiquitous factor binding to site D5 belongs to the AP-1 (*jun*) family (Fig. 4 and 5B). Recent studies have shown that AP-1 activity is modulated by the action of protein kinase C (1). As it is well known that PRL gene expression is stimulated by the hypothalamic releasing hormone TRH, which acts via Ca^{2+} and protein kinase C (22, 30, 35), it will be interesting to find out whether the D5 site plays a role in this TRH regulation.

The D8 site contains the perfect 10-bp palindromic sequence TGTTGCAACA. We have no knowledge of such a sequence having been reported as a protein-binding site. This suggests that the D8 sequence binds a novel *trans*-acting factor.

Comparison of the human and rat PRL genes. In the rat PRL gene, only two positive regulatory regions have been identified. Their respective locations are within the 400 bp of the proximal promoter sequences and in the distal 5′ region from −1530 to −1800 (36). These two regions are active only in pituitary-derived GH3 cells and contain four Pit-1-binding sequences (Fig. 1). We show here that these two regulatory regions are highly conserved in the rat and human PRL genes (Fig. 1, boxes I and III). They correspond to the proximal and distal regions of the human PRL gene.

It is possible that an additional regulatory region lies upstream from coordinate −3000 in the rat PRL gene as in the human gene, since Nelson et al. have not sought regulatory regions beyond this point (36, 37). This hypothesis is supported by an earlier report that proteins specifically bind to sequences located far upstream in the rat PRL gene, notably between positions −2000 and −4800 (45).

In addition to the three positive regulatory regions described above, our transfection assays suggest the presence, in the human PRL gene, of a slight negative region between positions −3500 and −1750 (Fig. 2). It is noteworthy that sequences in the rat PRL gene in the vicinity of coordinates −1800 to −2000 exhibit a similar negative effect (36, 37). That PRL gene expression might be subject to negative control has recently been proposed by Crenshaw et al., whose experiments with transgenic mice indicate that the distal enhancer (−1800 to −1530) and the proximal region (−422 to +33) are sufficient for the pituitary expression of the rat PRL gene but sequences flanking the distal enhancer are necessary to restrict expression to the correct cell type within the pituitary (8). As the human negative regulatory region comprises the D8, D9, and D10 binding sites (Fig. 3A), it will be interesting to investigate by mutational analysis whether these sequences play a role in this negative control. Furthermore, the D8 site is recognized by a ubiquitous factor whose concentration (or binding activity) is much lower in PRL-expressing GH3B6 cells than in PRL-nonexpressing cells (e.g., HeLa, BeWo, rat liver). It is thus tempting to speculate that this factor could act as a repressor. Experiments are under way to verify this hypothesis.

ACKNOWLEDGMENTS

We thank G. Schütz for pBLCAT2.

B.P. held a doctoral fellowship from FNRS (Belgium). M.L.V. and M.B. held a doctoral fellowship from IRSIA. This work was partly supported by a grant from SPPS (Belgium, Sciences de la vie, BIO/15) and the FNRS (Belgium, FRSM 3.4577.86).

LITERATURE CITED

1. **Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin.** 1987. Phorbol ester-inducible genes contain a common *cis* element recognized by a TPA-modulated *trans*-acting factor. Cell **49**:729–739.
2. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27**:299–308.
3. **Baumruker, T., R. Sturm, and W. Herr.** 1988. OBP100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. Genes Dev. **2**:1400–1413.
4. **Bodner, M., J.-L. Castrillo, L. E. Theill, T. Deerinck, M. Ellisman, and M. Karin.** 1988. The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell **55**: 505–518.
5. **Bohmann, D., T. J. Bos, A. Admon, T. Nishimura, P. K. Vogt, and R. Tjian.** 1987. Human proto-oncogene c-*jun* encodes a DNA-binding protein with structural and functional properties of transcription factor AP-1. Science **238**:1386–1392.
6. **Cao, Z., E. A. Barron, A. J. Carrillo, and Z. D. Sharp.** 1987. Reconstitution of cell-type-specific transcription of the rat prolactin gene in vitro. Mol. Cell. Biol. **7**:3402–3408.
7. **Cao, Z., E. A. Barron, and Z. D. Sharp.** 1988. Prolactin upstream factor I mediates cell-specific transcription. Mol. Cell. Biol. **8**:5432–5438.
8. **Crenshaw, E. B., K. Kall, D. M. Simmons, L. W. Swanson, and M. G. Rosenfeld.** 1989. Cell-specific expression of the prolactin gene in transgenic mice is controlled by synergistic interactions between promoter and enhancer elements. Genes Dev. **3**:959–972.
9. **Devereux, J., P. Haererli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
10. **Dolle, P., J.-L. Castrillo, L. E. Theill, T. Deerinck, M. Ellisman, and M. Karin.** 1990. Expression of GHF-1 protein in mouse pituitaries correlates both temporally and spatially with the onset of growth hormone gene activity. Cell **60**:809–820.
11. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316**:774–778.
12. **Fried, M. G., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interaction by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
13. **Galas, D. J., and A. Schmitz.** 1978. DNase footprinting, a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3170.
14. **Gorman, M. C., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79**:6777–6781.
15. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
16. **Guttierrez-Hartmann, A., S. Siddiqui, and S. Loukin.** 1987. Selective transcription and DNase I protection of the rat prolactin gene by GH3 pituitary cell-free extracts. Proc. Natl. Acad. Sci. USA **84**:5211–5215.
17. **Ingraham, H. A., R. Chen, H. J. Mangalam, H. P. Elsholtz, S. E. Flynn, C. R. Lin, D. M. Simmons, L. Swanson, and M. G. Rosenfeld.** 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell **55**:519–529.
18. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58**:799–839.
19. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct

transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42**:559–572.
20. **Keech, C. A., and A. Gutierrez-Hartmann.** 1989. Analysis of rat prolactin promoter sequences that mediate pituitary-specific and 3′,5′-cyclic adenosine monophosphate-regulated gene expression in vivo. Mol. Endocrinol. **3**:832–839.
21. **Landschulz, A. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
22. **Laverriere, J.-N., A. Tixier-Vidal, N. Buisson, A. Morin, J. A. Martial, and D. Gourdji.** 1988. Preferential role of calcium in the regulation of prolactin gene transcription by thyrotropin-releasing hormone in GH3 pituitary cells. Endocrinology **122**:333–340.
23. **Lefevre, C., M. Imagawa, S. Dana, J. Grindlay, M. Bodner, and M. Karin.** 1987. Tissue-specific expression of the human growth hormone gene is conferred in part by the binding of a specific *trans*-acting factor. EMBO J. **6**:971–981.
24. **Lemaigre, F. P., B. Peers, D. A. Lafontaine, M. Mathy-Hartert, G. G. Rousseau, A. Belayew, and J. A. Martial.** 1989. Pituitary-specific facor binding to the human prolactin, growth hormone, and placental lactogen genes. DNA **8**:149–159.
25. **Luckow, B., and G. Schütz.** 1987. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. **15**:5490.
26. **Lufkin, T., and C. Bancroft.** 1987. Identification by cell fusion of gene sequences that interact with positive *trans*-acting factors. Science **237**:283–286.
27. **Mangalam, H. J., V. R. Albert, H. A. Ingraham, M. Kapiloff, L. Wilson, C. Nelson, H. Elsholtz, and M. G. Rosenfeld.** 1989. A pituitary POU domain protein, Pit-1, activates both growth hormone and prolactin promoters transcriptionally. Genes Dev. **3**:946–958.
28. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**: 1237–1245.
29. **Manley, J. L.** 1984. Transcription of eukaryotic genes in a whole cell extract, p. 71–87. *In* B. D. Hames and S. J. Higgins (ed.), Transcription and translation—a practical approach. IRL Press, Oxford.
30. **Martin, T. F. J., and J. A. Kowalchyk.** 1984. Evidence for the role of calcium and diacylglycerol as dual second messengers in thyrotropin-releasing hormone action. Endocrinology **115**:1517–1535.
31. **Maurer, R. A.** 1985. Selective binding of the estradiol receptor to a region at least one kilobase upstream from the rat prolactin gene. DNA **4**:1–9.
32. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
33. **Muller, M., E. Bellefroid, C. Lambert, A. Belayew, and J. A. Martial.** 1986. Use of a bacterial, β-lactamase gene for studies of transcription regulation in eukaryotic cells. Arch. Int. Physiol. Biochem. **94**:B35.
34. **Müller, M. M., S. Ruppert, W. Schaffner, and P. Matthias.** 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. Nature (London) **336**:544–551.
35. **Murdoch, G. H., R. Franco, R. M. Evans, and M. G. Rosenfeld.** 1983. Thyrotropin-releasin hormone rapidly stimulates both transcription of the prolactin gene and the phosphorylation of a specific nuclear protein. J. Biol. Chem. **258**:15329–15335.
36. **Nelson, C., V. R. Albert, H. P. Elsholtz, L. I.-W. Lu, and M. G. Rosenfeld.** 1988. Activation of cell-specific expression of rat growth hormone and prolactin genes by a common transcription factor. Science **239**:1400–1405.
37. **Nelson, C., E. B. Crenshaw, R. Franco, S. A. Lira, V. R. Albert, R. M. Evans, and M. G. Rosenfeld.** 1986. Discrete *cis*-active genomic sequences dictate the pituitary cell type-specific expression of rat prolactin and growth hormone genes. Nature (London) **322**:557–562.
38. **O'Collaghan, C., A. Morris, S. A. Kirby, and A. H. Shingler.** 1972. Novel method for detection of β-lactamases by using a chromogenic cephalosporin substrate. Antimicrob. Agents Chemother. **1**:283–288.
39. **Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawakami, C. Balmaceda, R. A. Currie, and R. G. Roeder.** 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. Nature (London) **336**: 551–557.
40. **Schuster, W. A., M. N. Treacy, and F. Martin.** 1988. Tissue-specific *trans*-acting factor interaction with proximal rat prolactin gene promoter sequences. EMBO J. **7**:1721–1733.
41. **Serfling, E., M. Jasin, and W. Schaffner.** 1985. Enhancers and eukaryotic gene transcription. Trends Genet. **1**:224–230.
42. **Sharp, Z. D., S. Helsel, Z. Cao, E. A. Barron, and Y. Sanchez.** 1989. DNA recognition element required for PUF-1 mediated cell-type-specific transcription of the rat prolactin gene. Nucleic Acids Res. **17**:2705–2722.
43. **Theill, L. E., J.-L. Castrillo, D. Wu, and M. Karin.** 1989. Dissection of functional domains of the pituitary-specific transcription factor GHF-1. Nature (London) **342**:945–948.
44. **Truong, A. T., C. Duez, A. Belayew, A. Renard, R. Pictet, G. I. Bell, and J. A. Martial.** 1984. Isolation and characterization of the human prolactin gene. EMBO J. **3**:429–437.
45. **White, B. A., G. M. Preston, T. C. Lufkin, and C. Bancroft.** 1985. Detection of two chromatin proteins which bind specifically to the 5′-flanking region of the rat prolactin gene. Mol. Cell. Biol. **5**:2967–2974.