Three Sequence-Specific DNA-Protein Complexes Are Formed with the Same Promoter Element Essential for Expression of the Rat Somatostatin Gene†

OURANIA M. ANDRISANI, DAVID A. POT, ZHENG ZHU, AND JACK E. DIXON*

Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907

We identified three sequence-specific DNA-protein complexes which are formed after in vitro binding of nuclear extracts, derived from neuronal (CA-77, rat brain) or non-neuronal (HeLa) cells, to positions -70 to -29 of the rat somatostatin promoter. The protein(s) responsible for the formation of the three sequence-specific complexes was fractionated from rat brain whole cell extracts by DEAE-Sepharose chromatography. The critical contact residues of the factor(s) in each complex, as determined by methylation interference analyses, are located within positions -59 to -35, which is protected from DNase I digestion; these include the G residues of a TGACGTCA consensus also found in the cAMP-responsive human enkephalin (positions -105 to -76) and E1A-inducible adenovirus type 5 E_3 (positions -72 to -42) promoters. Competition assays with these heterologous promoters reveal that the factor(s) of each complex displays approximately 50-fold greater affinity for the somatostatin promoter-binding site. Synthetic oligonucleotides spanning positions -70 to -29 of the somatostatin promoter and containing single-base substitutions of the G residues in the TGACGTCA consensus were utilized in competition assays. The G residues located in the center of the module are the most critical determinants in the formation of the three sequence-specific complexes. Deletions disrupting the TGACGTCA consensus abolish not only formation of the three complexes in vitro but also expression of the somatostatin promoter in vivo, suggesting that formation of one or more of these complexes is essential for transcription of the rat somatostatin gene.

Regulation of transcription involves specific interactions of *trans*-acting factors with *cis*-acting DNA sequences. Molecular genetics has helped define *cis*-acting DNA elements for a number of viral and cellular eucaryotic promoters (reviewed in references 13 and 24). More recently, proteins that specifically bind to these *cis*-acting DNA elements have been identified by in vitro DNA-binding assays (4, 5, 17–19, 35). Accumulating evidence points to the existence of arrays of 8 to 12 base-pair recognition elements which are modular in nature. These elements contribute to the function of the promoter through their interaction with specific transcription factors to confer tissue-specific (36, 37, 41) temporal regulation (21, 31, 44) or activation in response to extracellular inducers (16, 23, 34).

We have focused our attention on the expression and regulation of genes which are selectively expressed in neuronal and endocrine cells (1, 8, 9, 12, 15, 27, 28, 42). In this report we have utilized in vitro DNA-binding assays to detect the factor(s) which interact in a sequence-specific manner with the promoter element required for expression of the neuropeptide hormone gene somatostatin. The somatostatin gene encodes the 14-residue peptide somatostatin, which is selectively expressed in the brain (32) as well as in endocrine cells of the pancreas (20) and gastrointestinal tract (33).

We have utilized the somatostatin-producing CA-77 cell line, which originated from a rat medullary thyroid carcinoma (2), to study the expression of the somatostatin gene. Expression studies in our laboratory (1) have demonstrated that only sequences downstream of position -60 are re-

quired for tissue-specific expression of this promoter in CA-77 cells, whereas a deletion to position -43 abolishes promoter function. The somatostatin promoter and its 5' deletions are not expressed in HeLa cells, which do not express the endogenous somatostatin gene. Expression studies of somatostatin promoter-chloramphenicol acetyltransferase (CAT) fusions in PC12 cells suggested that the somatostatin promoter is responsive to cyclic AMP (cAMP) (30). The region of the promoter required for cAMP responsiveness has been mapped between positions -60 and -29 and contains a cAMP consensus sequence, TGACGTCA, beginning at position -48. This is the same region that we found to be responsible for the tissue-specific expression of the somatostatin gene in CA-77 cells. The TGACGTCA consensus is found in most genes that appear to be regulated by cAMP, such as the rat phosphoenolpyruvate carboxykinase (45), human enkephalin (6), and the α-subunit of the human chorionic gonadotropin hormone (39). It is also found in the promoters of the early E1A-responsive genes of adenovirus (21, 22, 40, 43). Deletion analyses have confirmed the importance of the regions containing the TGACGTCA consensus in the expression of these promoters (for a review of adenovirus promoters, see reference 3).

We report here the identification of three sequence-specific DNA-protein complexes formed with the somatostatin promoter TGACGTCA consensus sequence. Deletions disrupting the TGACGTCA consensus disrupt not only the formation of the three complexes in vitro but also the expression of the promoter in vivo. Interestingly, the factor(s) forming these complexes interact with other promoters harboring the TGACGTCA consensus but display approximately 50-fold greater affinity for the somatostatin promoter-binding site.

MATERIALS AND METHODS

Gel retardation assays. Gel retardation analyses were carried out essentially as described by Carthew et al. (5) with a radiolabeled BglII-DraI (positions -70 to -29) 42-base-pair fragment of the rat somatostatin promoter. The BglII-DraI fragment was end labeled by the large fragment of DNA polymerase I or T_4 polynucleotide kinase and purified by gel electrophoresis on 20% acrylamide gels. Reactions were incubated in a volume of 20 μl at room temperature for 30 min. The binding assay was performed in 10 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid) (pH 7.9)–50 mM KCl–1 mM EDTA–5 mM dithiothreitol–5 mM $MgCl_2$–1 μg of poly(dI-dC)–10% glycerol–10 μg of extract–0.15 ng of DNA (15,000 cpm), and the reaction mixture was loaded onto a 4% polyacrylamide gel (acrylamide/bisacrylamide ratio, 80:1). Electrophoresis was at 10 V/cm in 6.7 mM Tris hydrochloride (pH 7.9)–3.3 mM sodium acetate–1 mM EDTA with buffer recirculation. The gels were then fixed in 10% acetic acid–10% methanol, dried, and autoradiographed.

DNase I footprinting of the somatostatin promoter fragment. The labeled probe for DNase I footprinting of the somatostatin promoter was prepared from plasmid pBxΔ-100 (1) by utilizing the HindIII-BamHI sites at positions -100 and +50, respectively. The noncoding strand was labeled by digesting and filling in plasmid pBxΔ-100 at the HindIII site (position -100), followed by a secondary BamHI digestion. The 150-base-pair labeled fragment was isolated by polyacrylamide gel electrophoresis.

A 2-ng (~20,000-cpm) sample of the end-labeled fragment was incubated with 50 μg of unfractionated nuclear CA-77 extract in a typical binding reaction for 30 min at room temperature and then subjected to DNase I (Bethesda Research Laboratories, Inc.) treatment for 1 to 5 min. After treatment with freshly diluted DNase I, the reaction was terminated by addition of 100 μl of 30 mM EDTA–1% sodium dodecyl sulfate–300 mM NaCl–250 μg of tRNA per ml. Samples were deproteinized by phenol-chloroform extraction, ethanol precipitated, and loaded onto an 8% denaturing polyacrylamide gel.

Methylation interference assay. For the methylation interference assay, the radiolabeled probe was prepared by end labeling the coding or the noncoding strand in separate reactions with T4 polynucleotide kinase. Annealing with the respective nonlabeled complementary strand was performed, and the double-stranded DNA was purified by electrophoresis on 20% polyacrylamide gels. Labeled probe (100 ng) was methylated by dimethylsulfate for 15 min (26) and then incubated with 200 μg of CA-77 nuclear extract in a large-scale binding reaction for 30 min at room temperature. The sample was loaded onto a preparative DNA-binding gel. The bands of interest were excised after autoradiography, and the radiolabeled DNA was isolated by electrophoresis onto NA-45 paper (Schleicher & Schuell Co.). The resulting radiolabeled DNA was cleaved with piperidine and loaded onto a 20% denaturing polyacrylamide gel.

Cell lines and extracts. HeLa cells were maintained in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal calf serum. The CA-77 cell line originated from a rat medullary thyroid carcinoma and expresses the endogenous somatostatin gene (2). It was maintained in defined medium, composed of 1:1 DMEM–F-10, supplemented with 5 μg of transferrin per ml, 10 μg of insulin per ml, and 3×10^{-8} M selenous acid. During passage of the

CA-77 cells, the plating medium was composed of 1:1 DMEM–F-10 and was supplemented with 5% horse serum and 2.5% fetal calf serum.

Whole cell extracts from HeLa, CA-77, and rat brain cells were prepared by the method of Manley et al. (25); nuclear and cytoplasmic extracts were prepared by the method of Dignam et al. (10). Protein concentrations were determined by the Bradford assay (3a) with serum albumin standards, and extracts were stored at -80°C.

DEAE-Sepharose chromatography. A DEAE-Sepharose column (90-ml column volume) was equilibrated with 10 mM HEPES (pH 7.9) containing 1 mM dithiothreitol, 1 mM EDTA, 5 mM $MgCl_2$, and 30 mM KCl. Rat brain whole cell extract (400 mg) prepared from 50 adult rats was dialyzed overnight against the same buffer and loaded onto the column; the flow rate was 30 to 35 ml/h. After the first A_{280} peak was washed off the column, a linear KCl gradient ranging from 30 to 300 mM was applied. Finally, the column was washed with buffer containing 500 mM KCl. Fractions (3.5 ml each) were measured by A_{280}, and DNA-binding activity was assayed by the gel retardation assay.

Recombinant plasmids and CA-77 cell transfections. The binding site deletions (Δ4 and Δ7) were constructed by digestion of plasmid pBxSST · t (1) at the unique AatII restriction site, starting at position -47, followed by S1 nuclease treatment and ligation. The size of the deletions was determined by sequencing. Positions -750 to +50 of the somatostatin gene containing the binding site deletions (Δ4 or Δ7) were cloned in front of the CAT gene as described earlier (1).

Plasmid DNA was transfected into CA-77 cells by electroporation as described earlier (1). CAT activity was assayed by the method of Gorman et al. (14) with an additional heat and EDTA treatment of the extract to inactivate cellular deacetylating activity (7).

RESULTS

Detection of protein factor(s) that binds to the rat somatostatin promoter. To identify the factors involved in the expression of the rat somatostatin gene, we examined the fragment between positions -70 and -29 of the promoter for specific DNA-protein interactions by utilizing the gel retardation assay (5, 11). Positions -70 to -29 of the somatostatin gene are shown in Fig. 1A. This region contained three nucleotides of the TATAA box at its 3' end and harbored sequences which resembled a G+C-rich box starting at position -64. It also contained the sequence TGACGTCA, between positions -48 and -42, which is similar to the consensus sequence for genes that appear to be regulated via a cAMP pathway (6, 30, 39, 45).

Figure 1B shows a typical gel retardation assay utilizing the radiolabeled promoter fragment spanning positions -70 to -29. In the absence of extract (Fig. 1B, lane 1) the DNA fragment migrated as a discrete band at the bottom of the gel. Whole cell, nuclear, and cytoplasmic extracts from somatostatin-producing CA-77 cells were incubated with the promoter fragment spanning positions -70 to -29 (Fig. 1B, lanes 2 through 4). Four bands, b1 through b4, were seen with the whole cell extract (lane 2); band b2 was the most prominent. The proteins forming bands b1 and b2 were primarily localized in the nuclear fraction (lane 4). In contrast, the proteins responsible for complexes b3 and b4 were found in both the nuclear and cytoplasmic fractions of the CA-77 extract (lanes 3 and 4). Identical gel retardation profiles and subcellular distribution of the binding activities
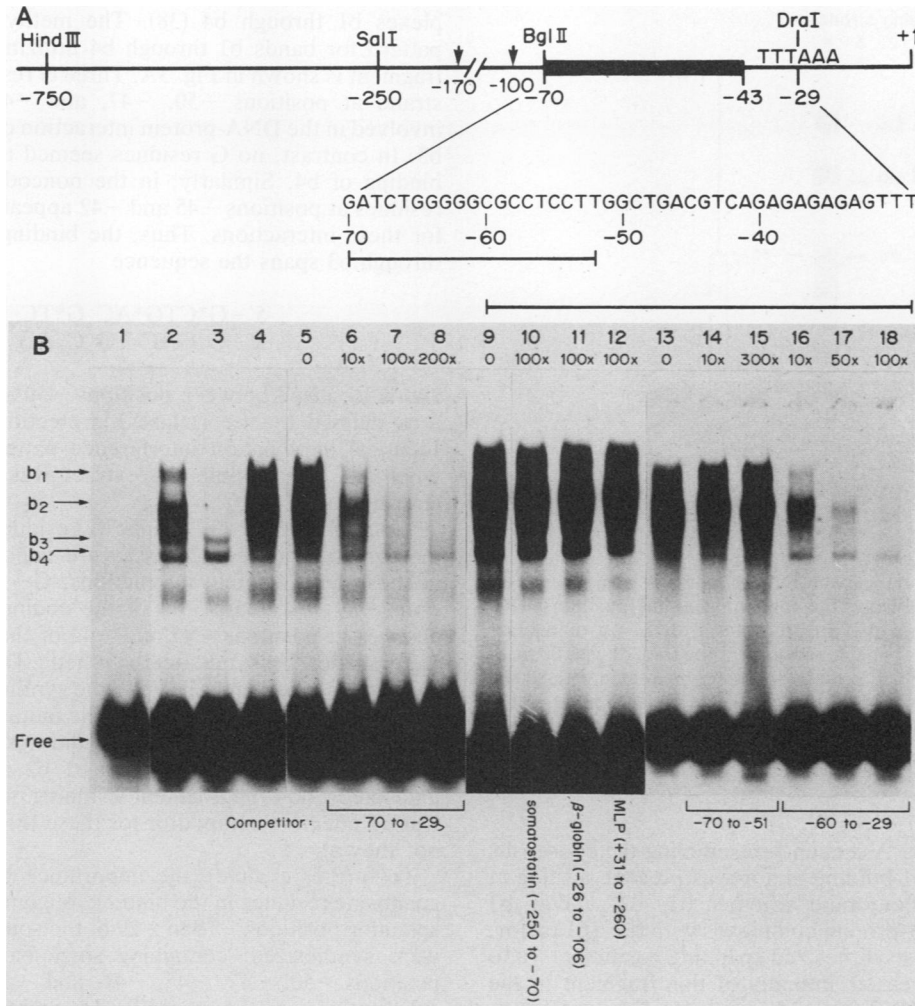
FIG. 1. Detection of protein factors that bind to the rat somatostatin promoter. (A) Schematic diagram and a limited restriction map of the rat somatostatin 5'-flanking region; the TATAA box homology shown is located at position −29; the CAAT box homology (not shown) is at position −93; a G+C-rich box consensus sequence is located starting at position −64. The blackened area between positions −70 and −43 indicates the region of the promoter required for expression of the gene in CA-77 cells. The coding-strand sequence of the BglII-DraI fragment used in gel retardation assays is shown below. Bars underneath this sequence indicate the smaller fragments utilized as competitor DNA in binding reactions. (B) Gel retardation profile of proteins binding to the fragment between positions −70 and −29 of the rat somatostatin promoter. Radiolabeled DNA (0.15 ng) was incubated with 10 µg of protein at room temperature before electrophoresis through a low-ionic-strength, 4% polyacrylamide gel. Fold molar excess of unlabeled competitor DNA in the binding reaction is indicated above each lane. Lanes: 1, free DNA fragment; 2, radiolabeled DNA fragment spanning positions −70 to −29 bound to CA-77 whole cell extract; 3 and 4, CA-77 S-100 and nuclear fractions, respectively; 5 through 8, competitions with increasing amounts of unlabeled somatostatin fragment spanning positions −70 to −29; 9 through 12, competitions with the upstream rat somatostatin promoter fragment (positions −250 to −70), rabbit β-globin (positions −106 to −26) or Ad5 major late promoter (positions −260 to +31); 13 through 15 and 16 through 18, competitions with increasing amounts of the rat somatostatin promoter fragments (positions −70 to −51 and −60 to −29, respectively).

were also observed with extracts from rat brain cells (data not shown). Proteinase K treatment of the extract before the binding reaction abolished the banding profile, indicating that discrete DNA-protein complexes are responsible for the retardation of the probe in the gel (data not shown).

To determine the sequence specificity of the DNA-protein interactions, competition assays were performed in the presence of increasing amounts of the unlabeled somatostatin promoter fragment (positions −70 to −29) (Fig. 1B, lanes 5 through 8). Bands b1, b2, and b3 showed complete competition at 100-fold molar excess of competitor (lanes 7 and 8), whereas b4 did not display any competition under these conditions, suggesting that the b4 complex is nonspecific. An adjacent fragment derived from the 5'-flanking region of the somatostatin gene (positions −250 to −70; Fig.

1A) did not compete with bands b1, b2, and b3 when included as a competitor in the binding reaction (Fig. 1B, lane 10). Competitions with other DNA sequences, such as the rabbit β-globin promoter (Fig. 1B, lane 11) or the adenovirus 5 major late promoter (Fig. 1B, lane 12) also did not affect the intensity of bands b1, b2, and b3. Likewise, a 10,000-fold excess by weight of poly(dI-dC) or a 750,000-fold molar excess of supercoiled pUC8 DNA did not affect the binding profile (data not shown). Thus, the site of interaction of specific nuclear factor(s) spans the region between positions −70 and −29 of the somatostatin promoter.

As an initial step in defining the somatostatin promoter region required for the interaction with the nuclear factor(s), specific portions of the region between positions −70 and −29 were synthesized and utilized as competitor DNA in the
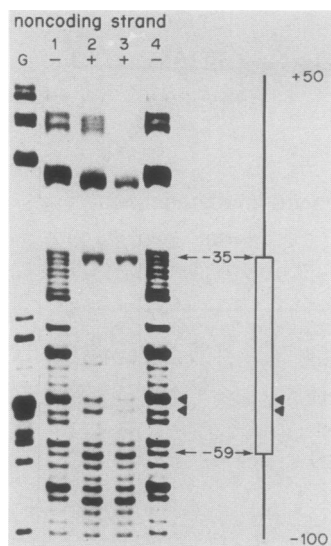
FIG. 2. Footprint analysis of the rat somatostatin promoter. DNase I protection analysis of the noncoding strand (positions −100 to +50) of the rat somatostatin promoter with 50 μg of nuclear CA-77 extracts. Lanes: 1 and 4, unbound probe treated with 80 ng of DNase I per ml for 2 min; 2, protein-bound probe treated with 800 ng of DNase I per ml for 5 min; 3, protein-bound probe treated with 8 μg of DNase I per ml for 5 min; G, Maxam and Gilbert G reaction. The protected region is noted to the right. The closed triangles indicate DNase I-hypersensitive sites.

gel retardation assay. A sequence resembling the consensus sequence for the Sp1-binding factor was present starting at position −64. To determine whether b1, b2, and/or b3 corresponds to DNA-protein complexes with the Sp1 factor, a DNA fragment was synthesized spanning positions −70 to −51 (Fig. 1A). Increased amounts of this fragment in the binding assay (Fig. 1B, lanes 13 through 15) did not displace any complex formed with the fragment spanning positions −70 to −29. This suggests that bands b1, b2, and b3 do not represent complexes with the Sp1-binding factor. In addition, a synthetic DNA fragment spanning residues −60 to −29 of the promoter lacking the G+C-rich box (Fig. 1B, lanes 16 through 18) competed effectively with bands b1, b2, and b3, indicating that the Sp1 factor is not involved in the generation of these bands.

DNase I footprinting of the rat somatostatin promoter fragment. We utilized the DNase I footprint assay with nuclear extracts from CA-77 cells to identify the region of the somatostatin promoter protected from DNase I digestion. The somatostatin promoter fragment (positions −100 to +50) labeled at the noncoding strand was incubated with nuclear CA-77 extracts and briefly treated with DNase I. The digestion products were analyzed on 8% denaturing polyacrylamide gels (Fig. 2). The protected region extended from positions −59 to −35 of the somatostatin promoter. Two strong DNase I-hypersensitive sites occurred at positions −55 and −54 in the protein-bound DNA. The footprint generated when the coding strand was labeled extended from position −59 to −38 (data not shown). At position −44 of the coding strand, a strong DNase I-hypersensitive site was seen with the protein-bound DNA.

Localization of the binding site. The methylation interference assay was performed to identify the G residues of the somatostatin promoter fragment (positions −70 to −29) that come into direct contact with the protein(s) forming com-

plexes b1 through b4 (38). The methylation interference pattern for bands b1 through b4 utilizing the somatostatin fragment is shown in Fig. 3A. Three G residues in the coding strand at positions −50, −47, and −44 appeared to be involved in the DNA-protein interaction of bands b1, b2, and b3. In contrast, no G residues seemed to be critical in the binding of b4. Similarly, in the noncoding strand, two G residues at positions −45 and −42 appeared to be important for these interactions. Thus, the binding site for bands b1 through b3 spans the sequence

$$5'-G*CTG*AC_*G*TC_*-3'$$
$$3'-C\ GAC\ TG^*C\ AG^*-5'$$

and is localized between positions −50 and −42, within the area defined by the DNase I protection assay (Fig. 3B). Identical methylation interference patterns were obtained when rat brain whole-cell extract was utilized (data not shown).

To confirm that the specific G residues identified by the methylation interference assay are important determinants in these DNA-protein interactions, G→A changes at positions −50, −47, and −44 of the coding strand and G→T changes at positions −45 and −42 of the noncoding strand were made. This mutated synthetic DNA fragment was identical in all other respects to the synthetic oligonucleotide spanning residues −60 to −29. The mutated fragment, when included in the binding reaction, did not compete even at a 300-fold molar excess with bands b1, b2, and b3, whereas the unlabeled wild-type fragment spanning positions −60 to −29 was an effective competitor for these three complexes (data not shown).

To further evaluate the importance of each of the five guanosine residues in the binding, oligonucleotide fragments spanning positions −70 to −29 of the somatostatin promoter were synthesized, containing single-base substitutions at positions −50, −47, −45, −44, and −42. The single-base substitutions are shown in Fig. 4A, and the oligonucleotides are referred to as point mutations (pm) 1 through 5. The competition assays performed in the presence of increasing amounts of each of the oligonucleotide fragments pm-1 through pm-5 are shown in Fig. 4B. Although from the methylation interference data the guanosine residue at position −50 seems to be involved in the interaction with the nuclear factor(s), the single-base substitution of G→T did not have any affect on the formation of the three sequence specific complexes. This suggests either that the G residue at position −50 is of secondary importance in the DNA-protein interactions or that the G→T change conserved the binding. The competition profile in the presence of increasing amounts of pm-2 and pm-5 DNA showed displacement of complexes b1 through b3 only at a 300-fold molar excess of competitor. In contrast, oligonucleotides pm-3 and pm-4 did not effectively compete for complex b1 through b3 formation against the wild-type fragment, even at a 300-fold molar excess of competitor DNA. Thus, the two most important guanosine residues are located within the center of the binding site at positions −45 and −44. The residues at positions −42 and −47 are also important determinants in the formation of complexes b1 through b3, although to a lesser degree. The least important seems to be the one at position −50.

Fractionation of the three somatostatin promoter-binding activities. Since whole cell rat brain extracts resulted in identical gel retardation and methylation interference profiles with somatostatin fragment spanning positions −70 to
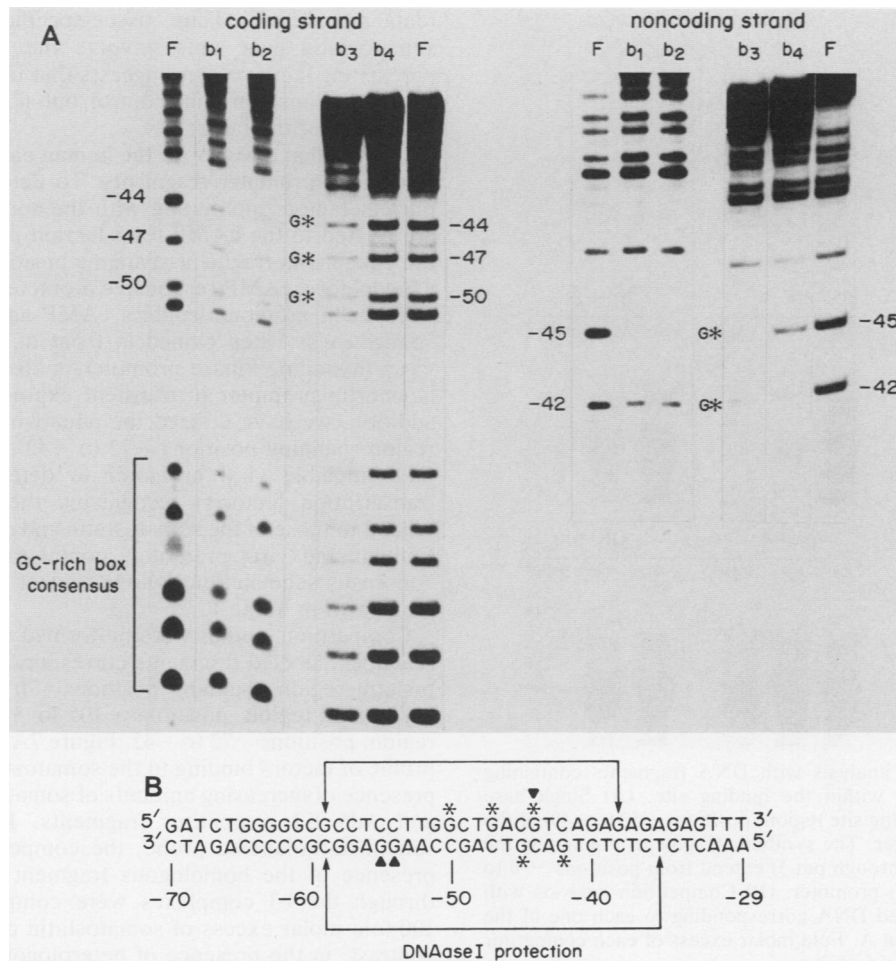
FIG. 3. Methylation interference analysis. (A) The end-labeled coding or noncoding strand of the somatostatin promoter fragment spanning positions −70 to −29 was methylated by treatment with dimethylsulfate for 15 min and incubated with CA-77 nuclear extract as described in the text. Reaction mixtures were electrophoresed on a preparative DNA-binding polyacrylamide gel. The radioactive fragments corresponding to bands b1 through b4 were isolated, cleaved with piperidine, and analyzed on a 20% sequencing gel. F, Maxam and Gilbert G reaction of the −70 to −29 DNA fragment. The nucleotide positions of apparent protein contacts are indicated. (B) Summary of DNase I protection and methylation interference analyses. The nucleotide sequence of the somatostatin promoter fragment between positions −70 and −29 is shown. The DNase I-protected area is indicated by arrows. Asterisks indicate the G residues that come into direct contact with factors forming complexes b1 through b3. Closed triangles indicate the DNase I-hypersensitive sites observed with the protein-bound DNA.

−29, we utilized rat brain extracts to fractionate the b1-through b3-binding activities by DEAE-Sepharose chromatography. The binding activities of b1 through b3 could be fractionated byDEAE-Sepharose chromatography by utilizing a linear KCl salt gradient (Fig. 5A). This fractionation step enabled us to study some characteristics of each of the binding activities. As an example (Fig. 5B), the gel retardation profile utilizing the b2 fractionated material resulted in formation of only the b2 complex, suggesting that a single binding activity does not produce the multiple binding activities seen in the nuclear extracts.

The properties of the fractionated b2 complex appear to be the same as those exhibited by the unfractionated nuclear extracts. This was shown by competition assays utilizing the somatostatin promoter fragments spanning positions –70 to –51 and −60 to –29. The fragment spanning positions −60 to –29 and containing the G→A and C→T changes (described above) was also included as a competitor in binding reactions with the DEAE-fractionated b2 fraction. Similarly to the CA-77 nuclear extract, only the wild-type fragment (positions −60 to −29) effectively competed for b2 binding when the b2-containing fraction was used (Fig. 5B).

Binding site deletions. To address the question of whether the binding of this set of nuclear factor(s) is important for the expression of the somatostatin promoter, we constructed two deletion mutations within positions −50 to −42 of the somatostatin promoter. One binding site deletion construct contained a 4-base-pair deletion of residues −46 to −43, and the second contained a 7-base-pair deletion of residues −47 to −41. Both deletion constructs (Δ4, Δ7) disrupted the TGACGTCA consensus starting at position −48. Somatostatin promoter fragments containing each binding site deletion were cloned in front of the CAT gene, resulting in plasmids pBxsstΔ4 and pBxsstΔ7. Deletion and control constructs were introduced into CA-77 cells by electroporation. After 48 h the cells were harvested and the extracts were assayed for CAT activity (Fig. 6A). Mutation pBxsstΔ4 retained only 9% of CAT activity when compared with the wild-type somatostatin promoter (positions −750 to +50), whereas the level of expression of mutation pBxsstΔ7 was similar to the expression level obtained with the inactive deletion pBxΔ-43. We conclude that the deleted sequences disrupting the TGACGTCA consensus are essential for expression of the somatostatin gene in CA-77 cells.
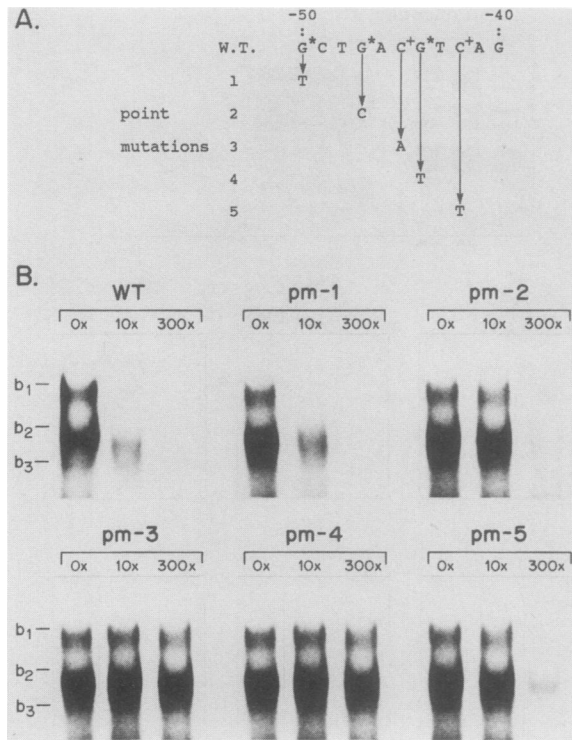
FIG. 4. Competition analysis with DNA fragments containing single-base substitution within the binding site. (A) Single-base substitutions in the binding site region (positions −50 to −40) of the rat somatostatin promoter. The synthetic fragments containing the point mutations (pm-1 through pm-5) extend from positions −70 to −29 of the somatostatin promoter. (B) Competition analysis with synthetic double-stranded DNA corresponding to each one of the point mutations shown in A. Fold molar excess of each competitor fragment is shown above each lane.

The radiolabeled fragment spanning positions −70 to −29 of the wild-type somatostatin promoter and the two deletion mutations were incubated either with nuclear CA-77 extract or the DEAE b2-containing fraction (Fig. 6B). The 4- and 7-base-pair deletion fragments exhibited reduced or no b1 through b3 binding (lanes 2, 3, 5, and 6) compared with the wild-type control (lanes 1 and 4). Interestingly, the nonspecific b4 complex was still present. In addition, competition assays to a level of 300-fold molar excess with the fragment spanning positions −70 to −29 obtained from the two deletion constructs did not result in any noticeable competition for any of the bands (data not shown). These studies demonstrate impaired binding of the nuclear factor(s) to the mutated somatostatin promoter and strongly suggest that formation of one or more of these DNA-protein complexes is required for expression of the somatostatin promoter in CA-77 cells.

Same binding activities in HeLa cells. One model for preferential expression of tissue-specific genes involves the existence of *trans*-acting factors only in the cell type in which the gene is expressed. To determine whether the somatostatin promoter-binding activities are distributed in a tissue specific manner, we examined extracts from HeLa cells for the existence of these binding activities. Surprisingly, although HeLa cells do not express somatostatin, HeLa cell extracts did produce gel retardation, competition, and subcellular distribution profiles of the binding activities that were identical to those observed with CA-77 extracts

(data not shown). Thus, tissue-specific expression of the somatostatin gene must involve the activation of these preexisting factors. This suggests that factor activation may alter the transcriptional control but not the DNA-binding properties of the factor(s).

Competition assays with the human enkephalin and adenovirus 5 E₃ promoter fragments. To determine whether the nuclear factors interacting with the somatostatin promoter are related to the cAMP transduction pathway, we utilized the enkephalin fragment spanning positions −105 to −76 as a model for a cAMP-responsive promoter. This region of the enkephalin promoter confers cAMP and phorbol ester responsiveness when cloned in front of the herpes simplex virus thymidine kinase promoter or the mouse proopiomelanocortin promoter in transient expression assays (6). In addition, we have utilized the adenovirus type 5 (Ad5) E3 region spanning positions −72 to −42 as an example of an E1A-inducible viral promoter to determine whether the transcription factor(s) recognizing the viral promoter is related to those of the somatostatin and/or enkephalin genes. Cellular and viral promoters containing the TGACGTCA consensus sequence in regions critical for their expression are shown in Table 1.

Competition studies were performed utilizing, in separate reactions, labeled fragments corresponding to the rat somatostatin region spanning positions −70 to −29, the human enkephalin region, positions −105 to −76, and the Ad5 E3 region, positions −72 to −42. Figure 7A shows the retention profile of factors binding to the somatostatin fragment in the presence of increasing amounts of somatostatin, enkephalin, and Ad5 E3 competitor fragments. In the case of the somatostatin-labeled probe, the competition profile in the presence of the homologous fragment shows that the b1 through the b3 complexes were completely displaced at 100-fold molar excess of somatostatin competitor DNA. In contrast, in the presence of heterologous competitor DNA (enkephalin or Ad5 E3) the complexes b1 through b3 were not completely displaced, even at 500-fold molar excess. Thus, the complexes formed with the somatostatin promoter have an apparent greater affinity for the somatostatin-binding region than for the enkephalin and Ad5 E3 regions.

To confirm the above observations, we performed the reciprocal experiments (Fig. 7B and C) utilizing the labeled enkephalin (positions −105 to −76) or the Ad5 E3 (positions −72 to −42) fragments as probes in the presence of increasing amounts of competitor DNA. The retention profiles (Fig. 7B and C) utilizing the labeled enkephalin or Ad5 E3 fragments displayed two major complexes (indicated by asterisk and arrow). The slowest-migrating complex (asterisk) corresponds to the binding activity forming complex b2 with the somatostatin promoter fragment at positions −70 to −29. This was confirmed by in vitro binding reactions utilizing each one of the three probes with the b2-containing DEAE-fractionated extract (unpublished observations). Interestingly, in binding reaction with the enkephalin- or Ad5 E3-labeled probes (Fig. 7B and C) the complex indicated by the asterisk was completely displaced at a 10-fold excess of somatostatin competitor DNA but not at a 10-fold excess of enkephalin or Ad5 E3 DNA. This observation confirms that the b2-forming activity displays higher affinity for the somatostatin binding site. Thus, although the factor(s) forming the three sequence-specific complexes with the somatostatin region at positions −70 to −29 can also bind to other TGACGTCA consensus-containing promoters, their binding affinities are not comparable.

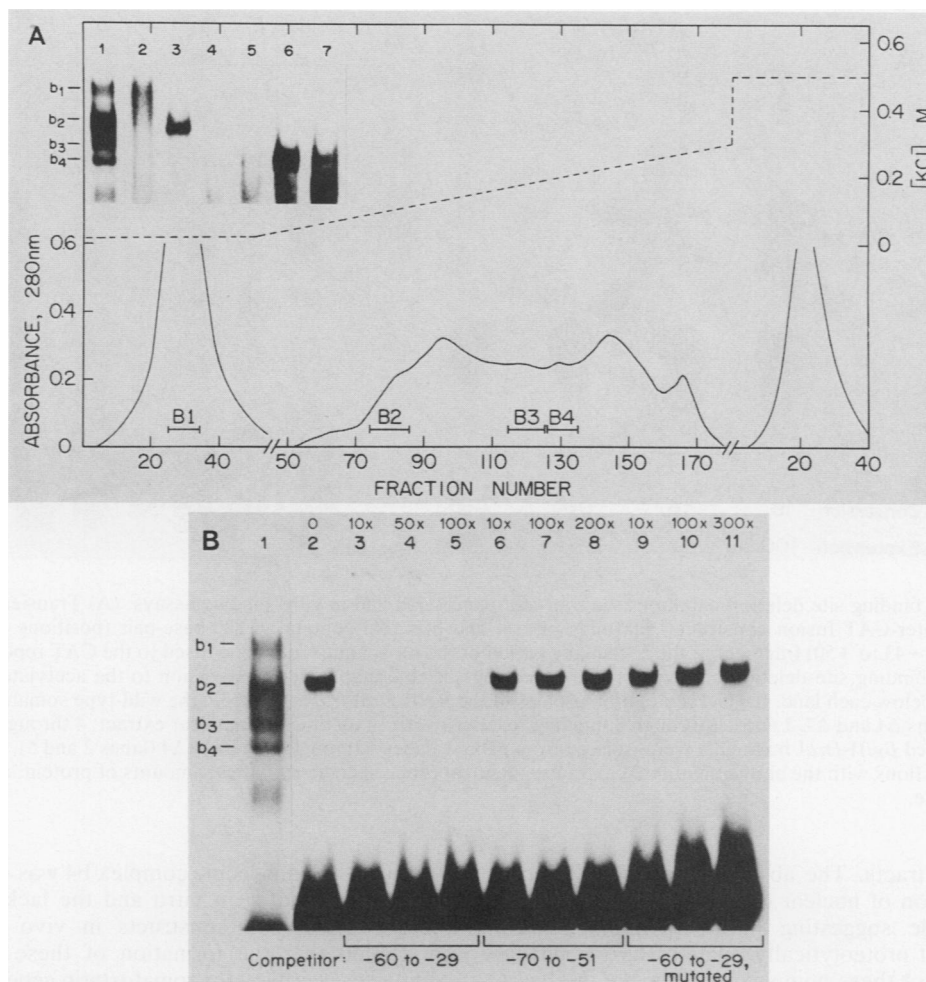Fortuitously, these reciprocal experiments (Fig. 7B and C)

FIG. 5. Fractionation of the binding activities of b1 through b3 by DEAE-Sepharose column chromatography. (A) A 400-mg sample of rat brain whole cell extract loaded onto a DEAE-Sepharose column (90 ml) and eluted with a linear KCl salt gradient, as shown in the profile. The binding activities of b1 through b4 were localized by a gel retardation assay. Inset lanes: 1, activity in the whole cell extract; 2, fractions 25 through 35; 3, fractions 75 through 85; 4 and 5, fractions 86 through 114; 6 and 7, fractions 115 through 125 and 125 through 136, respectively. (B) Properties of the b2-containing fraction. A 0.15-ng sample of radiolabeled somatostatin promoter fragment spanning positions −70 to −29 was incubated with 5 μg of protein from the b2-containing fraction. Lanes: 1, whole cell rat brain extract; 2 through 11, the b2-containing fraction in the absence of competitor (lane 2) and with increasing amounts of the following competitors; 3 through 5, fragment spanning positions −60 to −29; 6 through 8, fragment spanning positions −70 to −51; 9 through 11, mutated DNA fragment spanning positions −60 to −29. Fold molar excess of the respective unlabeled competitor DNA is indicated above each lane.

also identified an enkephalin-Ad5 E3-specific DNA-protein complex (shown by arrow, Fig. 7B and C). This DNA-protein complex did not exhibit detectable competition even at a 500-fold molar excess of somatostatin competitor DNA but competed with either enkephalin or Ad5 E3 fragments. This indicates that there appears to be a specific DNA-protein complex associated with the enkephalin-Ad5 E3 sequence.

DISCUSSION

We report here the identification of three DNA-protein complexes formed in a sequence-specific manner when nuclear CA-77, HeLa, or rat brain whole cell extracts interact with the region of the rat somatostatin promoter required for expression in CA-77 cells. The DNA-binding site is located within positions −59 to −35 of the somatostatin promoter and includes the TGACGTCA consensus, found in cAMP-responsive cellular promoters and E1A-inducible adenoviral promoters (Table 1). The protein-binding activities forming the sequence-specific complexes have been fractionated by DEAE-Sepharose chromatography from rat brain whole cell extracts. The b2-forming activity retains binding properties identical to those in the unfractionated nuclear extract. The nature of each complex is not known at present, but they all recognize and interact with the identical G residues within positions −50 to −42 of the somatostatin promoter. A possible explanation for these three sequence-specific complexes is that a single factor binds to the DNA but additional protein-protein interactions with the DNA-bound factor result in the formation of these multiple complexes. These complexes may also represent distinct posttranslational modifications of the same protein, or they may be distinct proteins, since they can be fractionated by ion-exchange chromatography. Alternatively, each factor may be the result of proteolytic cleavage. However, various protease inhibitors such as phenylmethylsulfonyl fluoride, leupeptin, and tolylsulfonyl phenylalanyl chloromethyl ketone were included in the preparation of CA-77 and
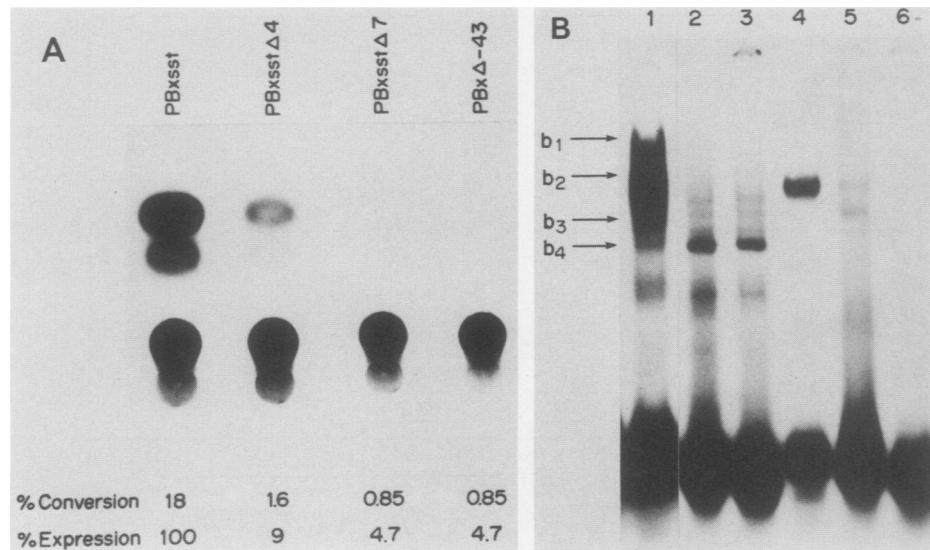
FIG. 6. Analysis of binding site deletion mutations via transient expression and in vitro binding assays. (A) Transient expression assays of somatostatin promoter-CAT fusion constructs. Plasmids pBxsst and pBxΔ-43 contain an 800-base-pair (positions −250 to +50) and a 93-base-pair (positions −43 to +50) fragment of the 5′-flanking region of the rat somatostatin gene fused to the CAT reporter gene. pBxsstΔ4 and pBxsstΔ7 are the binding site deletion constructs. The percentage of chloramphenicol conversion to the acetylated form and percent expression are shown below each lane. (B) Gel retardation profiles of the BglII-DraI fragments from the wild-type somatostatin promoter and the binding site deletions Δ4 and Δ7. Lanes: 1 through 3, binding reactions with 10 μg of CA-77 nuclear extract; 4 through 6, 5 μg of rat brain b2 fraction. Radiolabeled BglII-DraI fragments were isolated from pBxsst (lanes 1 through 4), pBxsstΔ4 (lanes 2 and 5), and pBxsstΔ7 (lanes 3 and 6) in binding reactions with the above-mentioned extracts. Identical probe specific activities, amounts of protein, and incubation times were used for each lane.

HeLa whole cell extracts. The absence of these inhibitors during the preparation of nuclear extracts did not alter the gel retardation profile, suggesting that these binding activities are probably not proteolytically related. Further physical characterization of these complexes will make the basis for these differences apparent.

Two experiments strongly support the involvement of one or more of these sequence-specific DNA-binding activities in the transcriptional activation of the somatostatin gene. First, deletions of 4 and 7 base pairs within the defined binding site of the nuclear factor(s) reduced the expression potential of the somatostatin promoter in CA-77 cells to a level of 9% and basal level of expression, respectively. Second, the mutated promoter fragments exhibited reduced or absence of b1 through b3 band formation in gel retardation assays. In

contrast, the nonspecific complex b4 was still formed. Thus, the lack of binding in vitro and the lack of expression of these two deletion constructs in vivo provide a strong correlation that the formation of these sequence-specific complexes is critical for somatostatin gene transcription. We are currently performing in vitro transcription assays to establish unequivocally whether this relationship is causal. In addition, these experiments will help distinguish which of the three complexes is responsible for the transcriptional activation of the rat somatostatin promoter.

There are various explanations for the existence of these three sequence-specific complexes (b1 through b3). A number of cellular (cAMP-responsive) and viral (adenovirus E1A-inducible) promoters contain the TGACGTCA consensus sequence (Table 1). Perhaps the complexes (b1 through

TABLE 1. TGACGTCA consensus-containing promoters

Protein	Sequence[a]													Map positions
Rat somatostatin	T T G	G*	C	T G* A C+ G* T C+ A	G A									−53 to −39
hCG[b] α-subunit	A A A	A	T	T G A C G T C A	T G									−129 to −115
Enkephalin	G G G	C	C	T G C G T C A	G C									−97 to −84
PEPCK[c]	G C C	C	C	T T A C G T C A	G A									−96 to −82
Ad2 E2A (e)	T G	G A G	A T G A C G T A	G T										−82 to −68
Ad2/Ad5 E3	C C C	T	G	T G A C G A A A	G C									−49 to −63
Ad2/Ad5 E3	T T C A G	A	T G A C T A A	C T										−101 to −86
Ad2 E4	G G A	A	G	T G A C G T A T	C G									−168 to −154

[a] Asterisks and plus signs indicate the important G residues of the rat somatostatin promoter, identified by the methylation interference assay.
[b] hGC, Human chorionic gonadotropin.
[c] PEPCK, Phosphoenolpyruvate carboxykinase.

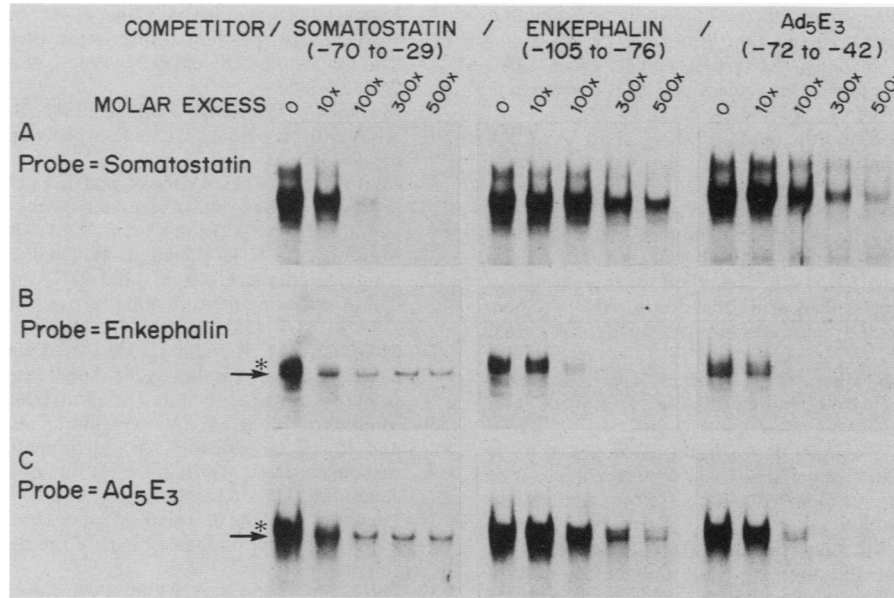FIG. 7. Comparison of the retention profile in the presence of rat somatostatin, human enkephalin, and Ad5 E3 promoter fragments. (A) Samples (10 µg) of CA-77 nuclear extracts were incubated with 0.15 ng of somatostatin (positions −70 to −29) radiolabeled probe in the presence of increasing amounts of unlabeled competitor. The molar excess of competitor is indicated above each lane. The competitor DNA included the following fragments: somatostatin (positions −70 to −29), enkephalin (positions −105 to −76), and Ad5 E3 (positions −72 to −42). The unlabeled competitor DNA was added 5 min before the addition of the labeled probe in the reaction mixture. Incubation was for 30 min at room temperature. (B and C) Identical conditions were utilized as described above for A, but the radiolabeled probe DNA was the human enkephalin (positions −105 to −76) and the Ad5 E3 (positions −72 to −42) fragments, respectively. Asterisk indicates the b2-binding activity. Arrow indicates the enkephalin-Ad5 E3-specific DNA-protein complex.

b3) detected by the in vitro binding assay with the somatostatin region spanning positions −70 to −29 represent not only complexes of somatostatin gene specific factors (functional binding) but also factors activating other TGACGTCA-containing promoters (nonfunctional binding). Alternatively, one or more of the three complexes may be necessary for tissue-specific expression of the rat somatostatin promoter. This agrees with the biology of the somatostatin gene, which is tissue specific in its mode of expression but responds to more widely acting inducers such as cAMP (30). Montminy and Bilezikjian (29) recently identified a nuclear protein in PC12 cells which selectively binds to positions −55 to −32 of the somatostatin gene. They showed that the affinity-purified factor can be phosphorylated in vitro by the catalytic subunit of the cAMP-dependent protein kinase. In addition, after stimulation of PC12 cells with forskolin a three- to fourfold increase in the phosphorylation of this protein occurs.

It is also of considerable interest that the factors b1 through b3 are present in cell lines which express somatostatin (CA-77) as well as in those that do not (HeLa). Sen and Baltimore (37) recently showed that NF-κB, a nuclear protein detected by in vitro binding assays only in cells that transcribe the immunoglobulin light-chain genes, is induced in the absence of protein synthesis by a posttranslational modification of a precursor factor found more widely in other cell types. Thus, a possible scenario to explain the selective expression of the somatostatin gene in CA-77 cells, but not HeLa cells, would involve tissue-specific activation of preexisting transcription factor(s) resulting in altered transcriptional activity without affecting its DNA-binding properties.

Our competition results with fragments from the somatostatin, enkephalin and Ad5 E3 promoters suggest that,

although the nuclear factor(s) forming the three sequence-specific complexes recognize the TGACGTCA module of these promoters, they display different binding affinities for the TGACGTCA element of the various genes. This premise would suggest a model of how genes expressed within the same cell and containing similar modular elements, such as the TGACGTCA element, could be differentially regulated by various inducers such as cAMP. The different binding affinities of trans-acting factor(s) can result in differential expression of two given genes containing similar modular elements.

ACKNOWLEDGMENTS

LITERATURE CITED

1. Andrisani, O. M., T. E. Hayes, B. Roos, and J. E. Dixon. 1987. Identification of the promoter sequences involved in the cell specific expression of the rat somatostatin gene. Nucleic Acids Res. 15:5715–5728.

2. Aron, D. C., M. Muszynski, R. Birnbaum, S. Sabo, and B. Roos. 1981. Somatostatin elaboration by monolayer cell cultures derived from transplantable rat medullary thyroid carcinoma: synergistic stimulatory effects of glucagon and calcium. Endocrinology 109:1830–1834.

3. Berk, A. J. 1986. Adenovirus promoters and E1A transactivation. Annu. Rev. Genet. 20:45–79.

3a. Bradford, M. M. 1976. A rapid and sensitive method for the

quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.

4. **Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science **234:**47–52.

5. **Carthew, R. W., L. A. Chodish, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43:**439–448.

6. **Comb, M., N. C. Birnberg, A. Seasholtz, E. Herbert, and H. M. Goodman.** 1986. A cyclic AMP- and phorbol ester-inducible DNA element. Nature (London) **323:**353–356.

7. **Crabb, D., and J. E. Dixon.** 1987. A method for increasing the sensitivity of chloramphenicol acetyltransferase (CAT) assays in extracts of transfected cultured cells. Anal. Biochem. **163:**88–92.

8. **Deschenes, R. J., R. S. Haun, C. L. Funckes, and J. E. Dixon.** 1985. A gene encoding rat cholecystokinin. J. Biol. Chem. **260:**1280–1286.

9. **Deschenes, R. J., L. J. Lorenz, R. S. Haun, B. A. Roos, K. J. Collier, and J. E. Dixon.** 1984. Cloning and sequence analysis of a cDNA encoding rat preprocholecystokinin. Proc. Natl. Acad. Sci. USA **81:**726–730.

10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.

11. **Fried, M., and D. M. Crothers.** 1981. Equilibria kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9:**6505–6525.

12. **Funckes, C. L., C. D. Minth, R. Deschenes, M. Magazin, M. A. Tavianini, M. Sheets, K. Collier, H. L. Weith, D. C. Aron, B. A. Roos, and J. E. Dixon.** 1983. Cloning and characterization of a mRNA-encoding rat preprosomatostatin. J. Biol. Chem. **258:**8781–8787.

13. **Gluzman, Y., and T. Shenk.** 1983. Enhancers and eukaryotic gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

14. **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.

15. **Hayes, T. E., and J. E. Dixon.** 1985. Z DNA in the rat somatostatin gene. J. Biol. Chem. **260:**8145–8156.

16. **Hayes, T. E., A. M. Kitchen, and B. H. Cochran.** 1987. Inducible binding of a factor to the c-*fos* regulatory region. Proc. Natl. Acad. Sci. USA **84:**1272–1276.

17. **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.

18. **Kadonaga, J. T., K. A. Jones, and R. Tjian.** 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. **11:**20–23.

19. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83:**5889–5893.

20. **Koerker, D. J., W. Ruch, E. Chideckel, J. Palmer, C. J. Goodner, J. Ensinck, and C. C. Gale.** 1974. Somatostatin: hypothalamic inhibitor of the endocrine pancreas. Science **184:**482–484.

21. **Kovesdi, I., R. Reichel, and J. R. Nevins.** 1986. Identification of a cellular transcription factor involved in E1A trans-activation. Cell **45:**219–228.

22. **Lee, K. A. W., and M. R. Green.** 1987. A cellular transcription factor E4F1 interacts with an E1A-inducible enhancer and mediates constitutive enhancer function in vitro. EMBO J. **6:**1345–1353.

23. **Lee, W., P. Mitchell, and R. Tjian.** 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell **49:**741–752.

24. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue specific gene expression. Science **236:**1237–1245.

25. **Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gefter.** 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc. Natl. Acad. Sci. USA **77:**3855–3859.

26. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.

27. **Minth, C. D., P. C. Andrews, and J. E. Dixon.** 1986. Characterization sequence, and expression of the cloned human neuropeptide Y gene. J. Biol. Chem. **261:**11974–11979.

28. **Minth, C. D., S. R. Bloom, J. M. Polak, and J. E. Dixon.** 1984. Cloning, characterization, and DNA sequence of a human cDNA encoding neuropeptide tyrosine. Proc. Natl. Acad. Sci. USA **81:**4577–4581.

29. **Montminy, M. R., and L. M. Bilezikjian.** 1987. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature (London) **328:**175–178.

30. **Montminy, M. R., K. A. Sevarino, J. A. Wagner, G. Mandel, and R. H. Goodman.** 1986. Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. Proc. Natl. Acad. Sci. USA **83:**6682–6686.

31. **Parker, C. S., and J. Topol.** 1984. A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an *hsp* 70 gene. Cell **37:**273–283.

32. **Patel, Y. C., and S. Reichlin.** 1978. Somatostatin in hypothalamus, extrahypothalamic brain, and peripheral tissues of the rat. Endocrinology **102:**523–530.

33. **Polak, J., A. G. E. Pearse, L. Grimelius, S. Bloom, and A. Arimura.** 1975. Growth-hormone release-inhibiting hormone in gastrointestinal and pancreatic D cells. Lancet **i:**1220–1222.

34. **Prywes, R., and R. G. Roeder.** 1986. Inducible binding of a factor to the c-*fos* enhancer. Cell **47:**777–784.

35. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell **43:**165–175.

36. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46:**705–716.

37. **Sen, R., and D. Baltimore.** 1986. Inducibility of κ-immunoglobulin enhancer-binding protein NF-κB by a posttranslational mechanism. Cell **47:**921–928.

38. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77:**122–126.

39. **Silver, B. J., J. A. Bokar, J. B. Virgin, E. A. Vallen, A. Milsted, and J. H. Nilson.** 1987. Cyclic AMP regulation of the human glycoprotein hormone α-subunit gene is mediated by an 18-base-pair element. Proc. Natl. Acad. Sci. USA **84:**2198–2202.

40. **SivaRaman, L., S. Subramanian, and B. Thimmappaya.** 1986. Identification of a factor in HeLa cells specific for an upstream transcriptional control sequence of an E1A-inducible adenovirus promoter and its relative abundance in infected and uninfected cells. Proc. Natl. Acad. Sci. USA **83:**5914–5918.

41. **Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore.** 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature (London) **323:**640–643.

42. **Tavianini, M. A., T. E. Hayes, M. D. Magazin, C. D. Minth, and J. E. Dixon.** 1984. Isolation, characterization, and DNA sequence of the rat somatostatin gene. J. Biol. Chem. **259:**11798–11803.

43. **Weeks, D. L., and N. C. Jones.** 1985. Adenovirus E3-early promoter: sequences required for activation by E1A. Nucleic Acids Res. **13:**5389–5402.

44. **Wiederrecht, G., D. J. Shuey, W. A. Kibbe, and C. S. Parker.** 1987. The *Saccharomyces* and *Drosophila* heat shock transcription factors are identical in size and DNA binding properties. Cell **48:**507–515.

45. **Wynshaw-Boris, A., T. Gross Lugo, J. M. Short, R. E. K. Fournier, and R. W. Hanson.** 1984. Identification of a cAMP regulatory region in the gene for rat cytosolic phosphoenolpyruvate carboxykinase (GTP). J. Biol. Chem. **259:**12161–12169.