MOLECULAR AND CELLULAR BIOLOGY, Oct. 1988, p. 4174–4184
0270-7306/88/104174-11$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 8, No. 10

Modulation of Epidermal Growth Factor Receptor Proto-Oncogene Transcription by a Promoter Site Sensitive to S1 Nuclease

ALFRED C. JOHNSON, YOSHIHIRO JINNO, AND GLENN T. MERLINO*

Laboratory of Molecular Biology, National Cancer Institute, 9000 Rockville Pike, 37/2D27, Bethesda, Maryland 20892

Received 23 February 1988/Accepted 24 June 1988

The epidermal growth factor (EGF) receptor is the functional target of the mitogen EGF and the cellular homolog of the avian erythroblastosis virus *erbB* oncogene product. Regulation of expression of the proto-oncogene encoding the EGF receptor can be elucidated by studying the structure and function of the gene promoter outside the confines of the cell. Previously, we reported the isolation of the human EGF receptor gene promoter. The promoter is highly GC rich, contains no TATA or CAAT box, and has multiple transcription start sites. An S1 nuclease-sensitive site has now been found 80 to 110 base pairs (bp) upstream from the major in vivo transcription initiation site. Two sets of direct repeat sequences were found in this area; both conform to the motif TCCTCCTCC. When deletion mutations were made in this region of the promoter by using either *Bal* 31 exonuclease or S1 nuclease, we found that in vivo activity dropped three- to fivefold, on the basis of transient-transfection analysis. Examination of nuclear protein binding to normal and mutated promoter DNAs by gel retardation analysis and DNase I footprinting revealed that two specific factors bind to the direct repeat region but cannot bind to the S1 nuclease-mutated promoter. One of the specific factors is the transcription factor Sp1. The results suggest that these nuclear *trans*-acting factors interact with the S1 nuclease-sensitive region of the EGF receptor gene promoter and either directly or indirectly stimulate transcription.

The epidermal growth factor (EGF) is a potent mitogen capable of stimulating protein and RNA synthesis, DNA replication, and, ultimately, cellular proliferation (5, 19, 30). It elicits these cellular responses by binding to the N-terminus of a 170-kilodalton cell surface glycoprotein, the EGF receptor. The receptor possesses a very active tyrosine kinase at its C-terminus and is capable of phosphorylating itself and other substrates upon EGF binding (4). By virtue of its extensive homology to the *erbB* oncogene product of the avian erythroblastosis virus, the EGF receptor is considered the cellular *erbB* proto-oncogene (7, 33, 47, 53). This proto-oncogene has been found to be overexpressed and occasionally amplified in a variety of malignant cell types (6, 27, 31, 37, 38, 54). More recently, we demonstrated that overexpression of the normal EGF receptor in NIH 3T3 cells by a retroviral vector induces malignancy and tumorigenicity (48).

The production of a protein so deeply involved in cellular proliferation and potentially responsible for malignant transformation must be carefully regulated in the normal cell. However, relatively little is known about control of EGF receptor gene expression. Previously, we reported the isolation, sequence, and partial characterization of the human EGF receptor gene promoter (23). The promoter has multiple transcription start sites and numerous protein-binding sites, including at least five Sp1-binding sites (24). The presence of multiple start sites may be due to the absence of an obvious TATA box (23). A chimeric plasmid containing the EGF receptor gene promoter ligated upstream of the chloramphenicol acetyltransferase (CAT) reporter gene (15) is active when transiently transfected into human KB epidermoid carcinoma or African green monkey kidney CV-1 cell lines (17, 23, 24).

* Corresponding author.

Promoter regions of active genes are frequently hypersensitive to the endonucleases DNase I and S1 (10, 29, 50). This sensitivity is thought to be the result of changes in the structure of the active chromatin, possibly because of the differential dissociation of nucleosomes and local alteration in DNA superhelicity (29, 50). Isolated eucaryotic promoter regions cloned into plasmid vectors maintain sensitivity to S1 nuclease, provided the plasmid is supercoiled (12, 14, 34, 35, 40, 44, 55).

Previously, we showed that the EGF receptor gene promoter chromatin is hypersensitive to DNase I in human A431 epidermoid carcinoma cells, which are known to overexpress the EGF receptor (13). Because the DNase I-hypersensitive sites are often characterized by sensitivity to the single-strand-specific endonuclease S1 (29, 50), we tested the cloned EGF receptor gene promoter for sensitivity to S1 nuclease. We report in this paper the identification of a site exquisitely sensitive to S1 nuclease residing in the vicinity of two pairs of direct-repeat sequences conforming to the motif TCCTCCTCC. We demonstrate that this site is necessary for optimal gene activity and is associated with nuclear proteins whose binding correlates with optimal transcription activity.

MATERIALS AND METHODS

Cell lines. The cultured cells used included a subclone of human epidermoid carcinoma KB cells (KB3-1), human epidermoid carcinoma A431 cells, and African green monkey kidney CV-1 cells. All cell lines were maintained in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum (GIBCO Laboratories).

Plasmid DNA manipulations. S1 nuclease and DNA ligase were purchased from Pharmacia. Most restriction endonucleases were obtained from New England Biolabs, Inc. The

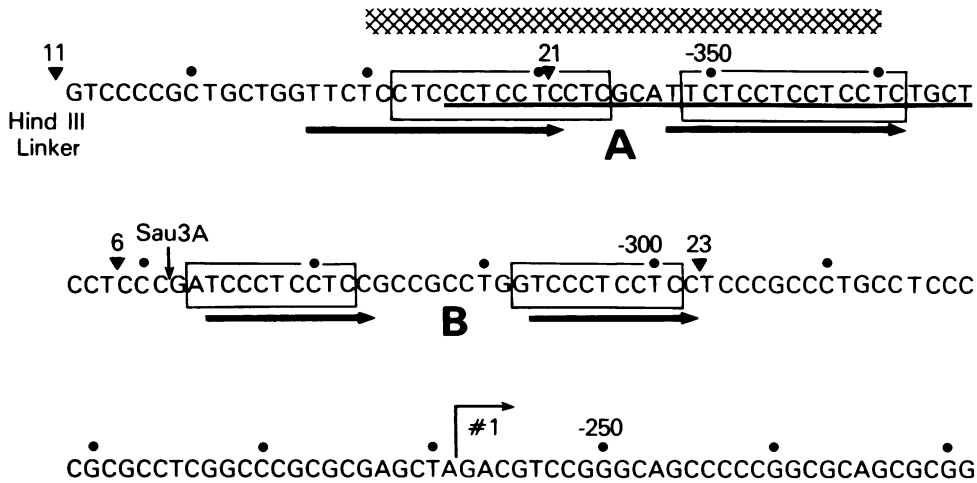


FIG. 1. EGF receptor gene promoter region. Sequences are numbered according to their positions relative to the translation start site (+1 at AUG). The major transcription start site (#1) is designated by the bent arrow at position −258. Heavy arrows underscore the location of the four sequences making up two sets of direct repeats (designated A and B). The cross-hatched box marks the general region sensitive to S1 nuclease. The underlined sequences are those deleted by S1 nuclease in the mutated promoter plasmid pS1d6. Boxed sequences designate putative Sp1-binding sites. Numbered arrowheads mark the upstream limit of *Bal* 31 digestion in cloned promoter mutations (−387 for 9H11, −359 for 9H21, −331 for 9H6, and −297 for 9H23). The numbering has been adjusted relative to the previously published promoter sequence (23) because of the addition of nucleotides at positions −282, −221, and −186 in that sequence.

simian virus 40 (SV40) enhancer region was obtained from pSV2CAT (15, 16) by using *Pvu*II and *Nco*I (converted to *Bam*HI with linker DNA). A 180-base-pair (bp) *Sal*I-*Sph*I fragment containing six GC boxes was obtained from the plasmid pdl41 (39).

To generate *Bal* 31 deletion mutants, a EGF receptor promoter-CAT construction containing a promoter fragment from −388 to −16 (pERCAT9; 24) was digested with *Nde*I (in the vector near the promoter insertion *Hin*dIII site), digested for various times with *Bal* 31, filled in with the Klenow fragment of DNA polymerase I, and ligated to *Xho*I linkers. DNA fragments containing the mutated promoter region, the CAT gene, and the RNA processing signals were isolated after digestion with *Apa*I. These DNAs were then ligated to the missing *Xho*I-*Apa*I vector fragment obtained from pERCAT2 (23), and the resulting complete plasmids were cloned into HB101 bacteria. One copy of the SV40 72-bp repeat enhancer sequence was then added into the *Bam*HI site at the end of the RNA polyadenylation signal.

To generate the S1 nuclease-mutated plasmids pS1d6 and pS1d7, a dimeric plasmid preparation of pGER9, consisting of the EGF receptor promoter fragment from −388 to −16, the CAT gene, and the RNA processing signals inserted into the *Hin*dIII and *Bam*HI sites of the plasmid pGEM3 (Promega), was digested with S1 nuclease (see below for a description of conditions). The linear monomeric form of the plasmid was then isolated from an agarose gel. The ends of these molecules were made blunt with Klenow fragment, religated, and cloned into N38 bacteria. One copy of the SV40 enhancer sequence was later inserted into the *Bam*HI site.

Deletion mutants were characterized by dideoxy DNA sequencing (43) by primer extension of RNA transcribed from the Sp6 promoter in the Gemini vector containing the mutated EGF receptor promoter DNA.

S1 nuclease sensitivity assay. S1 digestion of supercoiled plasmid DNA was performed as described elsewhere (29, 55). Briefly, DNA at a concentration of 0.1 μg/ml was digested in 30 mM sodium acetate (pH 4.5)–300 mM NaCl–0.2 mM EDTA–3 mM $ZnCl_2$ with 5 U of S1 nuclease per μg of DNA at 42°C for 25 min. At this time, more than half of the DNA was in a linear form. Digested DNA was extracted with phenol-chloroform-isoamyl alcohol (25:24:1) and ethanol precipitated. The S1 nuclease-sensitive site was then mapped by digesting S1 nuclease-cleaved DNA with a variety of restriction enzymes according to manufacturer specifications and electrophoresing on agarose or polyacrylamide gels.

DNA transfection and CAT assay. Supercoiled plasmid DNAs purified by double banding in CsCl gradients were transfected into either KB or CV-1 cells by calcium phosphate coprecipitation (15, 16, 22). Basically, 10 μg of DNA precipitated with calcium phosphate was added to cultured cells plated the day before at 4×10^5 or 8×10^5 cells per 100-mm-diameter dish for KB or CV-1 cells, respectively, and fed 3 h earlier with fresh complete media. After 3.5 to 4 h, the cells were glycerol shocked (30 s for KB cells and 3 min for CV-1 cells) and fed with complete media. After 48 h, extracts were prepared from the transfected cells, and CAT assays were performed by using acetyl-CoA and [^{14}C]chloramphenicol (15, 16). All assays were carried out with equal amounts of extract protein. Techniques were checked with a Rous sarcoma virus–β-galactosidase plasmid as an internal control.

RNA isolation and primer extension. Total RNA was isolated from monkey CV-1 cells (five 100-mm-diameter plates for each construction) at 48 h after plasmid DNA transfection by solubilization in guanidine isothiocyanate and centrifugation through a CsCl cushion (3). RNA concentrations were determined by UV absorbance and confirmed by formaldehyde-agarose gel electrophoresis. The latter technique was also used to check the quality of the RNA. A 15-μg sample of this RNA was hybridized to 0.1 pmol of a 5′-end-labeled CAT-specific DNA 24-mer in 40% formamide–0.4 M NaCl–1 mM EGTA [ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′,-tetraacetic acid]–40 mM PIPES

[piperazine *N*,*N*′-bis(ethanesulfonic acid); pH 6.4] at 42°C for 3 h. This primer, which hybridizes to the region between residues 4920 and 4943 of pSV2CAT (15, 16), was then extended by reverse transcriptase and analyzed by 8 M urea–polyacrylamide gel electrophoresis as described previously (36). Extended primers marking the RNA start sites were visualized by autoradiography.

DNA-protein binding assays. The gel retardation assay was performed by using a modification of the method of Singh et al. (45) as described elsewhere (24). Briefly, 0.5 to 1.0 ng of a DNA fragment 5′ end labeled with P-32 was incubated with 10 to 30 μg of A431 nuclear protein extract at room temperature for 20 min in 20 μl of 50 mM Tris hydrochloride (pH 7.5)–250 μg of bovine serum albumin per ml–5% glycerol–5 mM EDTA-2.8 μg of poly(dI-dC) · poly(dI-dC). DEAE-sepharose chromatography was used to create fractions BA and BC, as described elsewhere (26). The DNA-protein mixture was analyzed on a native 4% polyacrylamide gel (45). For some competition experiments, excess cold DNA fragments were preincubated with the nuclear protein extract at room temperature for 5 min before addition of the labeled fragment. Polyacrylamide gels were dried before autoradiography.

DNase I footprinting was performed as described elsewhere (8, 9). A 469-bp EGF receptor promoter *Mst*II-*Hin*dIII fragment P-32 5′ end labeled at the *Mst*II site was incubated with A431 nuclear extract previously fractionated by heparin-agarose chromatography, as described elsewhere (24). Routinely, protein fractions eluted from the heparin-agarose at NaCl concentrations of 0.3 and 0.6 M were incubated with end-labeled DNA, the mixture was digested with DNase I, and the resulting DNA was analyzed by 8 M urea–polyacrylamide gel electrophoresis and autoradiography. Affinity-purified Sp1 (2, 25) was generously provided by James Kadonaga and Robert Tjian (University of California, Berkeley, Calif.).

RESULTS

S1 nuclease-sensitive site in EGF receptor gene promoter. Upon examination of the EGF receptor gene promoter, two pairs of direct repeats were discovered (Fig. 1). One pair (designated B in Fig. 1), located between nucleotides −298 and −326 (relative to the translational start site), consisted of a perfect 10-bp repeat sequence (39 to 67 bp upstream of the major in vivo transcription initiation site). A second pair (designated A in Fig. 1), found between nucleotides −339 and −373, shared 14 of 15 bp (80 to 114 bp upstream of the major in vivo transcription start site). Both repeats had only pyrimidine residues on one strand and purine residues on the complement strand, and both conformed to the motif TCCTCCTCC. DNA containing stretches of (dC-dT) · (dA-dG), singly or repeatedly, have been shown to be associated with sensitivity to the endonuclease S1 both in vitro and in vivo (see Discussion).

To determine whether S1 nuclease sensitivity could be detected in this region, supercoiled plasmid DNA containing the EGF receptor gene promoter (−388 to −16) was used. When the promoter-containing plasmid pGER9 was sequentially cut with S1 nuclease and *Eco*RI, two new DNA fragments were generated that were not seen when *Eco*RI was used alone (Fig. 2, lane 3). Two new bands were also obtained when S1 nuclease digestion was followed by *Pvu*II cutting (Fig. 2, lane 4). This S1 nuclease-sensitive site was localized to the region between −340 and −370, coincident with the position of repeat A (Fig. 1). The same region was sensitive to the single-strand nuclease P1 as well (data not shown). It seems that sensitivity to S1 nuclease requires the supercoiled plasmid structure because *Eco*RI-restricted linear DNA was not digested by S1 nuclease (Fig. 3A, lane 4).

DNA region containing S1 nuclease-sensitive direct repeat is necessary for optimal EGF receptor gene expression. To test

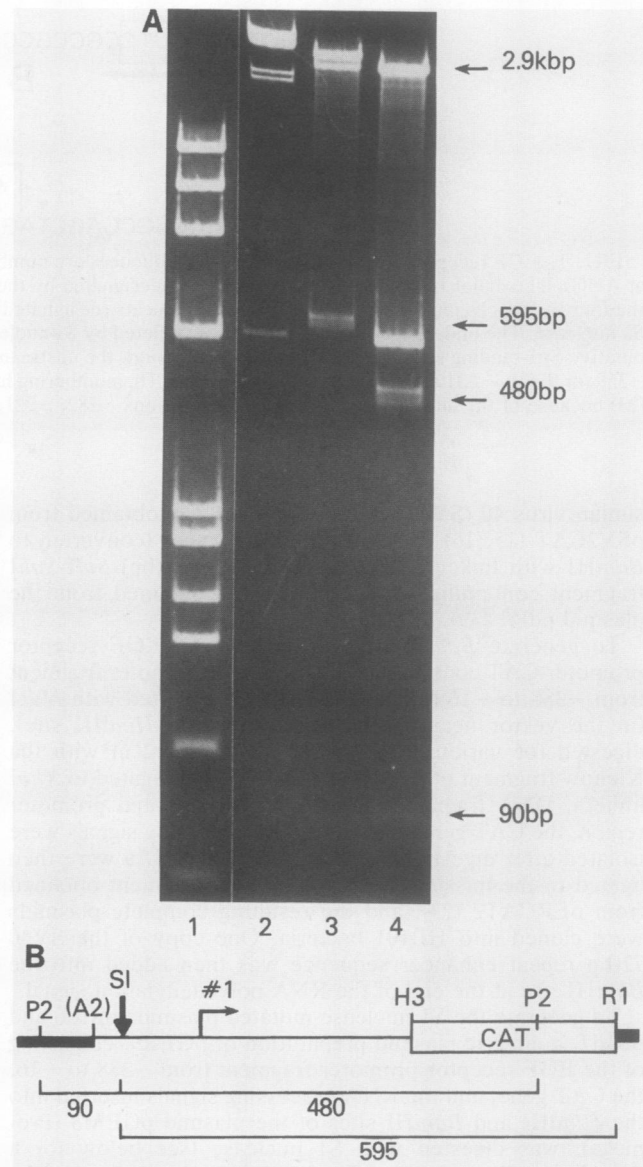


FIG. 2. Fine mapping of S1 nuclease-sensitive site in EGF receptor gene promoter. (A) Ethidium bromide-stained polyacrylamide gel. Lane 1, *Hae*III-digested φx marker DNA; lane 2, *Hin*dIII-digested λ marker DNA; lane 3, plasmid pGER9 digested sequentially with S1 nuclease and then with *Eco*RI, a procedure yielding two fragments of 2.9 kbp and 595 bp; lane 4, pGER9 digested with S1 nuclease and then with *Pvu*II, a procedure yielding two additional fragments of 480 and 90 bp. (B) Schematic representation of promoter region. The thick line is the vector DNA, and the thin line is promoter DNA. The box is the 250-bp *Hin*dIII-*Eco*RI CAT fragment. The bent arrow marks the major transcription start site (#1). The heavy vertical arrow indicates the S1 nuclease-sensitive site derived from this data. P2, *Pvu*II; H3, *Hin*dIII; R1, *Eco*RI; A2, *Ava*II (lost during cloning).

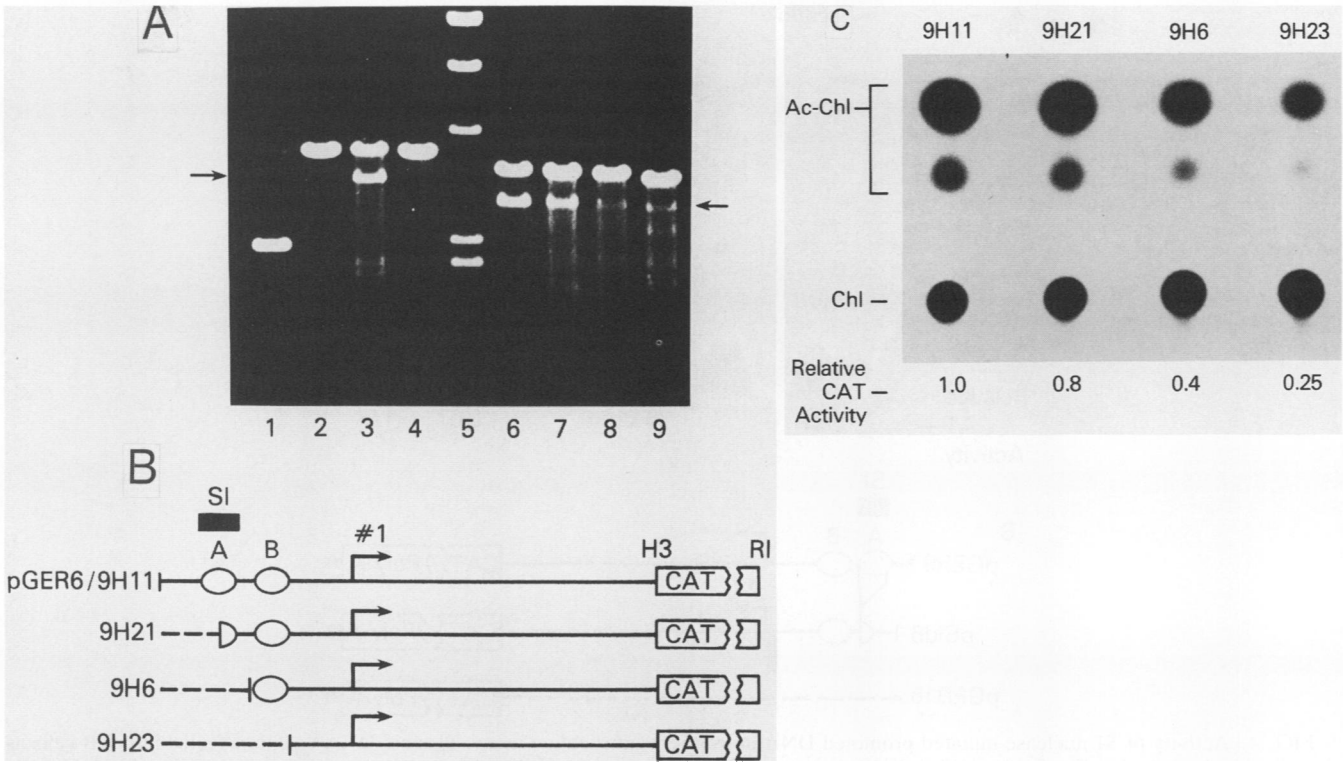


FIG. 3. Characterization of *Bal* 31 deletion mutations. (A) Ethidium bromide-stained agarose gel. Lanes 1 to 4, Supercoiled plasmid pGER6 was either untreated (lane 1), digested with *Eco*RI (RI) (lane 2), digested first with S1 nuclease and then with *Eco*RI (lane 3), or digested first with *Eco*RI and then with S1 nuclease (lane 4). Lane 5, *Hin*dIII-digested λ marker DNA. The left arrow shows one of the bands generated by S1 nuclease. The smaller second band was electrophoresed off the gel. Lanes 6 to 9, Digestion of *Bal* 31-generated promoter mutation plasmid 9H11 (lane 6), 9H21 (lane 7), 9H6 (lane 8), or 9H23 (lane 9) first with S1 nuclease and then with *Eco*RI. (B) Structures of *Bal* 31 mutations. The CAT *Hin*dIII-*Eco*RI (H3-RI) 250-bp fragment is boxed, and the major (#1) transcription start site is marked by the bent arrow. Each circle represents one pair of the direct repeats A and B (Fig. 1), and the black box marks the location of the S1 nuclease-sensitive region. Dashed lines indicate the extent of digestion by *Bal* 31. The vector for these plasmids was pGEM4. (C) Activity of *Bal* 31-mutated EGF receptor gene promoter DNA in a transient-transfection assay. Plasmid DNA (10 μg) was transfected into human KB cells by calcium phosphate coprecipitation. After 48 h, extracts were prepared and CAT activity was determined. These plasmids contain one copy of the 72-bp repeat SV40 enhancer region. CAT activity is presented relative to that of 9H11 (activity taken as 1.0). Chl, [^{14}C]chloramphenicol; Ac-Chl, acetylated [^{14}C] chloramphenicol.

whether the direct-repeat region is involved in EGF receptor gene expression, the promoter region, previously placed upstream of the bacterial CAT reporter gene, was digested with the exonuclease *Bal* 31. The resulting mutations eliminated different amounts of the two pairs of direct repeats (Fig. 1, arrowheads and Fig. 3B). As the sequences containing the repeats were eliminated, the S1-sensitive site disappeared (Fig. 3A, lanes 6 to 9). Each of the four mutated plasmids was separately transfected into human KB cells, and after 48 h, extracts were prepared for determination of CAT activity. As more of the repeat sequences were deleted, CAT activity continued to decrease (Fig. 3C). When both pairs of repeats were deleted (9H23), activity dropped by about fourfold. These results suggest that this DNA region is involved in EGF receptor gene regulation.

To more directly test the role of the S1-sensitive site, S1 nuclease was used to mutate this region in pGER9. After blunt-end religation, DNAs were cloned into HB101. Two cloned plasmids, pS1d6 and pS1d7, were isolated and sequenced. They were found to have 31- and 33-bp deletions in the upstream direct repeat A, respectively, a finding that confirms the location of the S1-sensitive site. One of these, pS1d6, was analyzed further. The parent plasmid pGER9 and the mutant pS1d6 are identical, except for the 31-bp S1-sensitive sequences; both contain the CAT gene and RNA processing signals of pSVOCAT and the vector pGEM3. In addition, another pair of plasmids was constructed that contained one copy each of the SV40 72-bp repeat enhancer element. A comparison of the CAT activity generated by transfection of these two promoters into cultured cells should reveal any change in gene transcription caused by the removal of repeat A.

The normal and S1-mutated plasmids with or without enhancers were transfected into monkey CV-1 cells. After 48 h, extracts were prepared from the cells for determination of CAT activity. The mutated promoter was 3.2-fold less active than the normal promoter when both contained a single copy of the SV40 enhancer and 2.1-fold less active without enhancers. Because CV-1 cells are monkey cells and make very low levels of EGF receptor, we used the normal and mutated plasmids (pGER9 and pS1d6) to transfect human KB epidermoid carcinoma cells, which contain about 100,000 receptors per cell (54). The KB cells are less active than CV-1 cells as transfection recipients, and we had to use enhancers in these plasmids to accurately compare CAT activities. The promoter with the S1 mutation was 4.7-fold less active than the normal promoter (Fig. 4), a result corroborating those obtained with CV-1 cells.

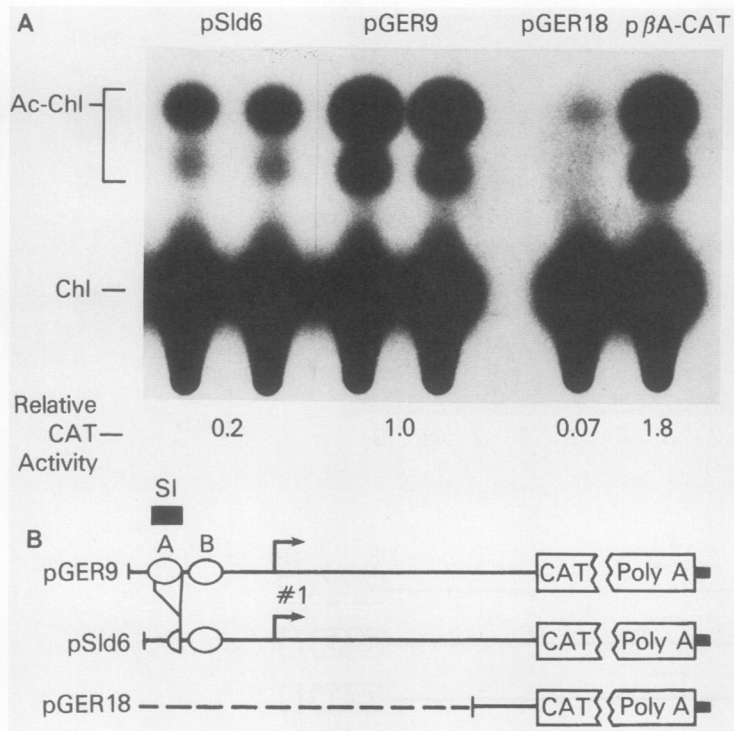
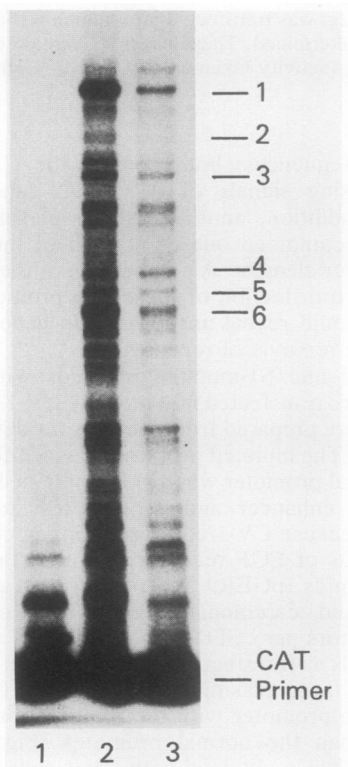


FIG. 4. Activity of S1 nuclease-mutated promoted DNA in transient-transfection assay. Plasmid DNA was transfected into KB cells as described in the legend to Fig. 3. (A) CAT activity is presented relative to that of pGER9 (activity taken as 1.0). pβA-CAT is a chicken β actin promoter-CAT construction which was included only as a positive control. The plasmid pGER18 (26) had no detectable activity and served as a negative control. (B) Schematic representation of the promoter S1 nuclease mutations. The open boxed regions depict the full CAT gene and RNA processing signals, and the open circles indicate the two sets of direct repeats A and B. The bent arrow indicates the #1 transcription start site, and the black box indicates the S1 nuclease-sensitive region. The dashed line indicates deleted DNA. The vector for pGER9 and pS1d6 was pGEM3, and for this experiment, both plasmids contained a single copy of the SV40 enhancer region.



These CAT assay results suggest that DNA sequences sensitive to nuclease S1 are involved in EGF receptor gene regulation. The CAT assays are indirect, albeit useful, measurements. To more directly ascertain the effects of this deletion, total RNAs were isolated from CV-1 cells transfected with either the normal pGER9 plasmid or the mutated pS1d6 plasmid. These RNAs were hybridized to a labeled CAT-specific 24-mer, and the primer was extended with reverse transcriptase. The EGF receptor gene promoter-driven CAT RNA was four- to fivefold higher in the normal promoter than in the mutated promoter (Fig. 5). No major

FIG. 5. Primer extension analysis of RNA isolated from monkey CV-1 cells transfected with normal (pGER9) and S1 nuclease-mutated (pS1d6) plasmid DNA. These plasmids contained one copy of the SV40 enhancer region. The 5′-end-labeled CAT-specific 24-mer used as primer for reverse transcriptase is shown at the bottom. Shown are the results of primer extensions of 10 μg of tRNA (lane 1), 15 μg of RNA isolated from CV-1 cells transfected with pGER9 (lane 2), and 15 μg of RNA isolated from CV-1 cells transfected with pS1d6 (lane 3). Numbers at the right indicate putative transcription start sites (23). The major start site is #1. Multiple start sites are presumably caused by the absence of a TATA box (Fig. 1). The high background is due to an elevated GC content in this region of the EGF receptor gene, which apparently causes premature termination of the transcriptase. However, the vast majority of the activity is specific for the EGF receptor gene promoter-CAT plasmid. Primer extension of untreated CV-1 RNA resulted in no substantial signal (data not shown).

changes in the transcription start sites were detected (Fig. 5). The RNA results were in complete agreement with the CAT protein data.

Binding of nuclear proteins to direct-repeat region of EGF receptor gene promoter. The S1-sensitive region may be involved in gene regulation by providing binding sites for nuclear *trans*-acting factors. If so, DNA-binding assays should show that proteins which can bind to the normal EGF receptor gene promoter cannot bind to the S1 nuclease-mutated promoter. To test this idea, a 70-bp *Hin*dIII-*Sau*3A fragment containing direct repeat A was isolated from the normal promoter, and a 39-bp *Hin*dIII-*Sau*3A fragment was isolated from the mutated promoter (Fig. 6 and 7). Both fragments were end labeled with P-32, separately incubated with nuclear extracts from A431 cells, and electrophoresed on a native polyacrylamide gel. Two prominent bands (Fig. 6A, lane 2, arrows 1 and 2) associated with the normal 70-bp promoter appeared in the gel. Their migration would suggest that they are DNA-protein complexes. Both bands represent specific binding because their intensity was substantially diminished when a 100-fold molar excess of unlabeled 90-bp fragment from the normal promoter (the 70-bp *Hin*dIII-*Sau*3A fragment cloned into pGEM3 and then cut with *Hin*dIII and *Eco*RI) was included in the DNA binding reaction (Fig. 6A, lane 3). An unlabeled 59-bp fragment from the pS1d6-mutated promoter (the 39-bp *Hin*dIII-*Sau*3A fragment cloned into pGEM3) competed poorly for binding (Fig. 6A, lanes 6 to 8); however, the 59-bp S1d6 fragment did compete at high concentrations for band 2. The latter was probably caused by the recreation of half of direct repeat A in the S1d6 mutation (TCCTCCCTCC; Fig. 1). Another unlabeled normal promoter fragment (the 215-bp *Hin*dIII-*Bgl*I fragment) was also used as a specific competitor (Fig. 7). Interestingly, even a 10-fold molar excess of this larger fragment, which contains both the upstream and downstream TCC repeat regions (repeats A and B, respectively) and two strong Sp1-binding sites (24), successfully competed with the labeled probe for binding to the unidentified protein(s) (Fig. 7a, lanes 1 to 5).

In contrast to the DNA fragment containing repeat A from the normal promoter, the S1 nuclease-mutated 39-bp DNA fragment was incapable of binding nuclear protein, as judged by the absence of slowly migrating bands in the gel retardation assay (Fig. 7b, lane 9).

Sp1 is one of two separate factors that bind to S1 nuclease-sensitive promoter regions. A clue to the identity of one of the nuclear DNA-binding proteins was revealed upon close examination of the DNA sequence of the two pairs of promoter direct repeats. Repeat B contained two 10-bp sequences that completely match the Sp1 consensus se-

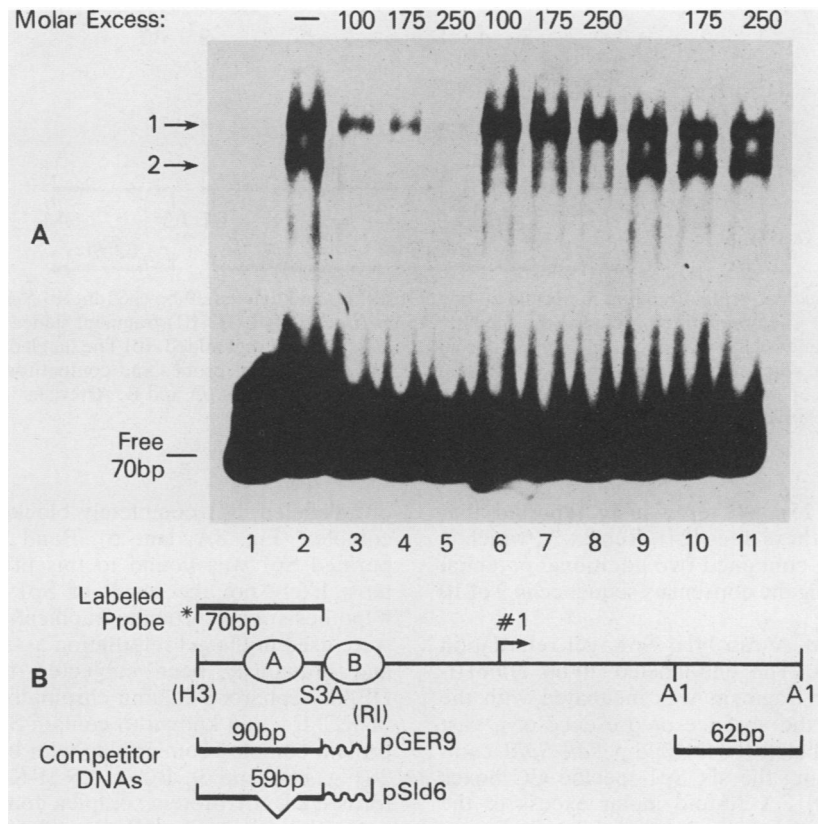


FIG. 6. Detection of specific EGF receptor gene promoter DNA-binding protein by gel retardation assay. (A) A431 nuclear extract was incubated with a 5'-end-labeled 70-bp *Hin*dIII-*Sau*3A (H3-S3A) fragment, and the resulting DNA-protein complexes were resolved by native polyacrylamide gel electrophoresis and autoradiography. Two retarded bands are visualized (lane 2, bands 1 and 2). Competitor DNAs included a 90-bp subcloned normal promoter fragment (from pGER9; lanes 3 to 5), a 59-bp subcloned S1-deleted promoter fragment (from pS1d6; lanes 6 to 8), and a nonspecific downstream 62-bp *Ava*I (A1) fragment (lanes 10 and 11). Lane 1, No extract was added; lane 9, 125 ng of additional poly(dI-dC) · poly(dI-dC) was included. (B) Schematic representation of the probe and competitors relative to the promoter. Wavy lines are vector sequences. RI, *Eco*RI (cloning site); #1, major transcription start site. The two open circles indicate the two sets of direct repeats A and B. Asterisks indicate the positions where the probes were 5' end labeled.

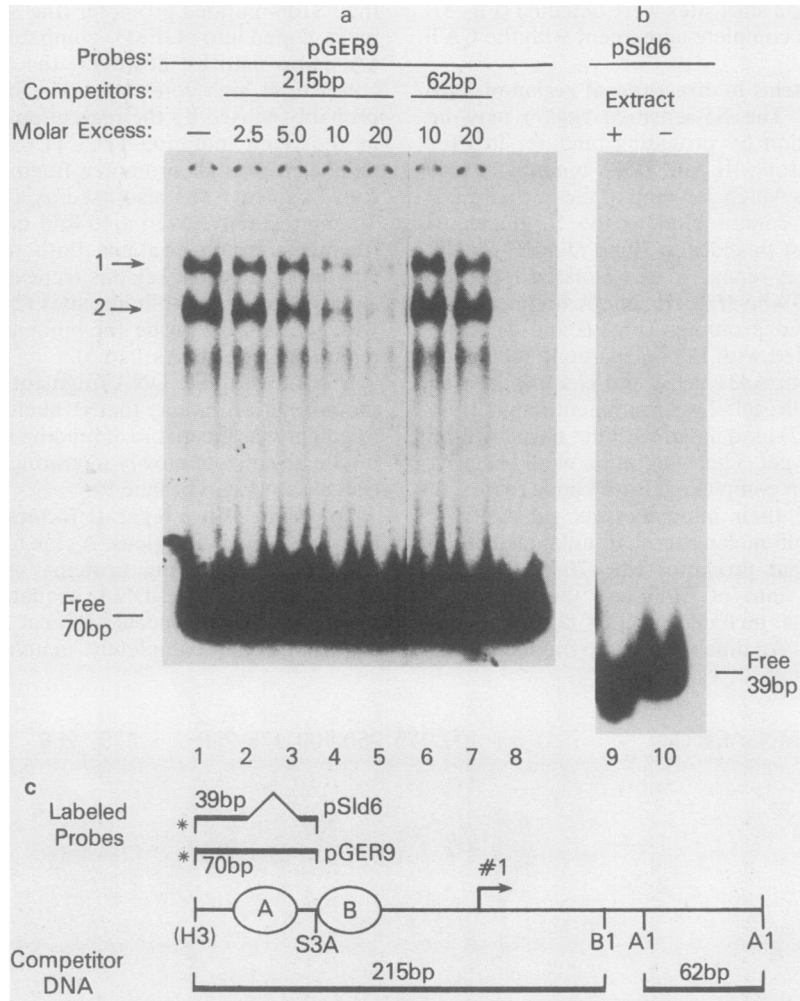


FIG. 7. Comparison of nuclear protein binding to normal 70-bp (pGER9) (a) and deleted 39-bp (pS1d6) (b) 5′-end-labeled promoter DNA by gel retardation assay. (a) The competitors included the specific 215-bp *Hin*dIII-*Bgl*I (H3-B1) fragment (lanes 2 to 5) and the nonspecific 62-bp *Ava*I (A1) fragment (lanes 6 and 7). Lane 1, No competitor added; lane 8, no extract added. (b) The labeled pS1d6 fragment was either incubated alone (lane 10) or with A431 extract (lane 9). (c) Schematic representation of probes and competitors. S3A, *Sau*3A; #1, major transcription start site. Two open circles mark the positions of the two sets of direct repeats A and B. Asterisks indicate the positions where the probes were 5′ end labeled.

quence-binding site (2, 25). We previously reported that purified Sp1 can bind to these sites (24). Repeat A, which is sensitive to S1 nuclease, contained two additional potential Sp1-binding sites matching the consensus sequence in 9 of 10 positions (Fig. 1).

To test whether repeat A can bind Sp1, gel retardation assays were again used. The end-labeled 70-bp *Hin*dIII-*Sau*3A normal promoter fragment was incubated with the A431 nuclear extract in the absence or presence of a vast molar excess of an unlabeled SV40 180-bp *Sal*I-*Sph*I competitor fragment containing the six Sp1-specific GC boxes (from plasmid pdl41 [39]). A 30-fold molar excess of the unlabeled dl41 fragment effectively competed for the formation of the complex designated band 1 (Fig. 8A, lane 3). Band 2 was not affected by the presence of the Sp1-binding fragment (Fig. 8A, lane 3).

Purified Sp1 was then incubated with the end-labeled 70-bp normal promoter fragment. Sp1 bound to the promoter DNA and formed a radioactive complex comigrating with band 1 (Fig. 8A, lane 4). Furthermore, a 30-fold molar excess of unlabeled dl41 completely blocked the formation of this complex (Fig. 8A, lane 5). Band 2 did not appear when purified Sp1 was bound to this promoter fragment and is most likely not the result of Sp1 binding. To prove this hypothesis, two chromatographically distinct A431 fractions were used in the gel retardation assay. A431 nuclear extract had previously been subjected to heparin-agarose and DEAE-sepharose column chromatography (26). BA, a 0.12 M-KCl fraction known to contain Sp1 (26), formed a DNA-protein complex comigrating with band 1 but not with band 2 (Fig. 8B, lane 7). BC, a 0.5 M-KCl fraction lacking Sp1, formed a DNA-protein complex comigrating with band 2 but not with band 1 (Fig. 8B, lane 8).

These gel retardation assay results indicate that one of the two factors capable of binding to the S1 nuclease-sensitive promoter region was Sp1. The other factor, which was not Sp1, was associated with the TC-rich repeat A region and is referred to as a TC factor or TCF.

The gel retardation assay is a very quick and sensitive technique for detecting DNA binding; however, it reveals

little information about the boundaries encompassed by the factor while bound to the DNA. We, therefore, employed the DNase I footprinting technique. First, purified Sp1 was incubated with a 5'-end-labeled 469-bp *Mst*II-*Hin*dIII fragment, and the resulting complexes were exposed to DNase I. Three footprints were recognizable (Fig. 9A, lane Sp1, boxes 1 to 3), corresponding to those previously identified in the EGF receptor gene promoter (24). Sp1 box 3 contained both of the Sp1-binding sites in repeat B (nucleotides −299 to −327). A much weaker footprint was associated with repeat A, probably because of the single-base-pair deviation from the Sp1 consensus sequence (Fig. 9A, lane Sp1).

The DNase I footprint assay was repeated by using A431 nuclear extracts. Footprints obtained with crude extracts were weak, and so a 0.6 M-NaCl fraction from a heparin-agarose column was employed; this 0.6 M fraction was known to contain both Sp1 and TCF (Fig. 8, lane 6). The three Sp1 footprints were again generated when A431 extracts were used (Fig. 9A, lane 0.6), and an additional footprint was located approximately between −340 and −370, corresponding to the TC-rich, S1 nuclease-sensitive repeat A. Furthermore, this DNase I footprint persisted when the 0.6 M A431 fraction was incubated in the presence of excess unlabeled dl41 (Fig. 9B); dl41 prevented Sp1 from binding to the labeled promoter DNA (Fig. 8). This result indicates that the footprint found between −340 and −370 is not Sp1. The DNase I footprint data support the gel retardation assay results in lending credence to the hypothesis that direct repeat A can bind multiple nuclear proteins.

DISCUSSION

The significance of the presence of DNase I and S1 nuclease-sensitive sites in eucaryotic DNA has been the subject of much experimentation and discussion. Generally, differences in sensitivity to nucleases have been associated with alterations in chromatin structure and consequent gene expression (10, 11, 52). Sensitivity to S1 nuclease has been studied most extensively in isolated supercoiled plasmid DNA and has been applied to native chromatin and isolated nuclei as well (29, 51). In vivo, DNase I- and S1 nuclease-sensitive sites will coexist at the same locus (29, 51).

Previously, we discovered the existence of a DNase I hypersensitive site in the 5'-flanking promoter region of the EGF receptor gene (23). We report in this paper the identification of an S1 nuclease-sensitive site residing in the vicinity of two pairs of novel direct repeats of the EGF receptor gene promoter cloned into plasmid vectors. This site resides approximately 100 bp upstream of the major in vivo transcription start site. All four elements in the two pairs of repeats consist of stretches of from 10 to 14 pyrimidine residues on one strand and conform to the same

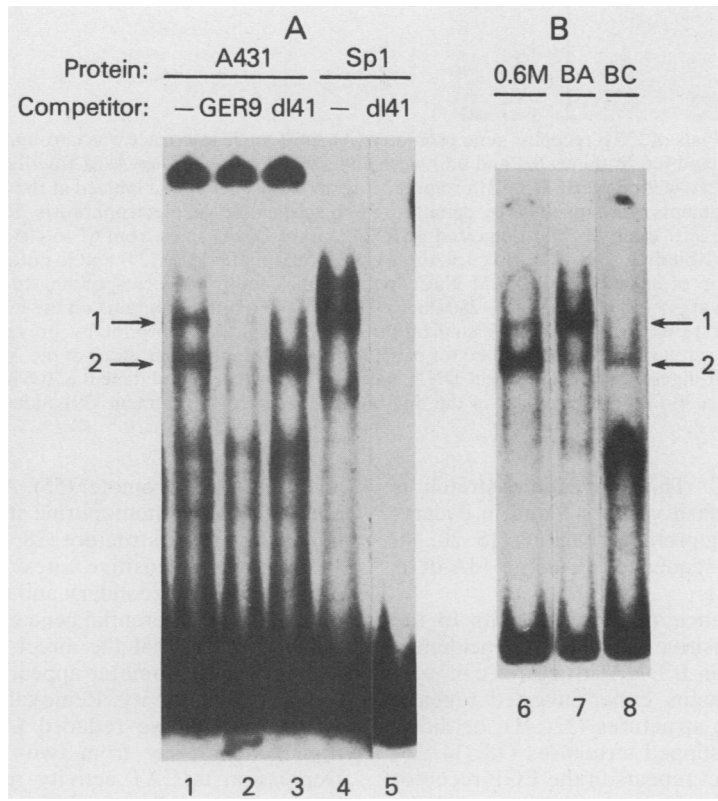


FIG. 8. One factor binding to direct repeat A of EGF receptor gene promoter is Sp1. The 70-bp *Hin*dIII-*Sau*3A probe is defined in Fig. 6. The gel retardation assay is described in the legend to Fig. 6. (A) Incubation of promoter probe with whole A431 nuclear extract (lane 1), with whole nuclear extract plus a 150-fold molar excess of the specific 215-bp *Hin*dIII-*Bgl*I promoter fragment (Fig. 7) (lane 2), with whole nuclear extract plus a 30-fold molar excess of a six-GC box-containing 180-bp *Sal*I-*Sph*I SV40 dl41 fragment (39) (lane 3), with 8 ng of purified Sp1 plus 100 ng of poly(dI-dC) · poly(dI-dC) (lane 4), and with 8 ng of Sp1 and a 30-fold molar excess (36 ng) of the 180-bp dl41 fragment (lane 5). (B) The gel retardation assay was performed as described by Singh et al. (45). Shown are results of the incubation of the 70-bp promoter probe with the 0.6 M NaCl heparin-agarose A431 fraction (lane 6), the DEAE-sepharose A431 fraction BA (lane 7), and the DEAE-sepharose A431 fraction BC (lane 8). Arrows 1 and 2 indicate the two major gel retardation bands.

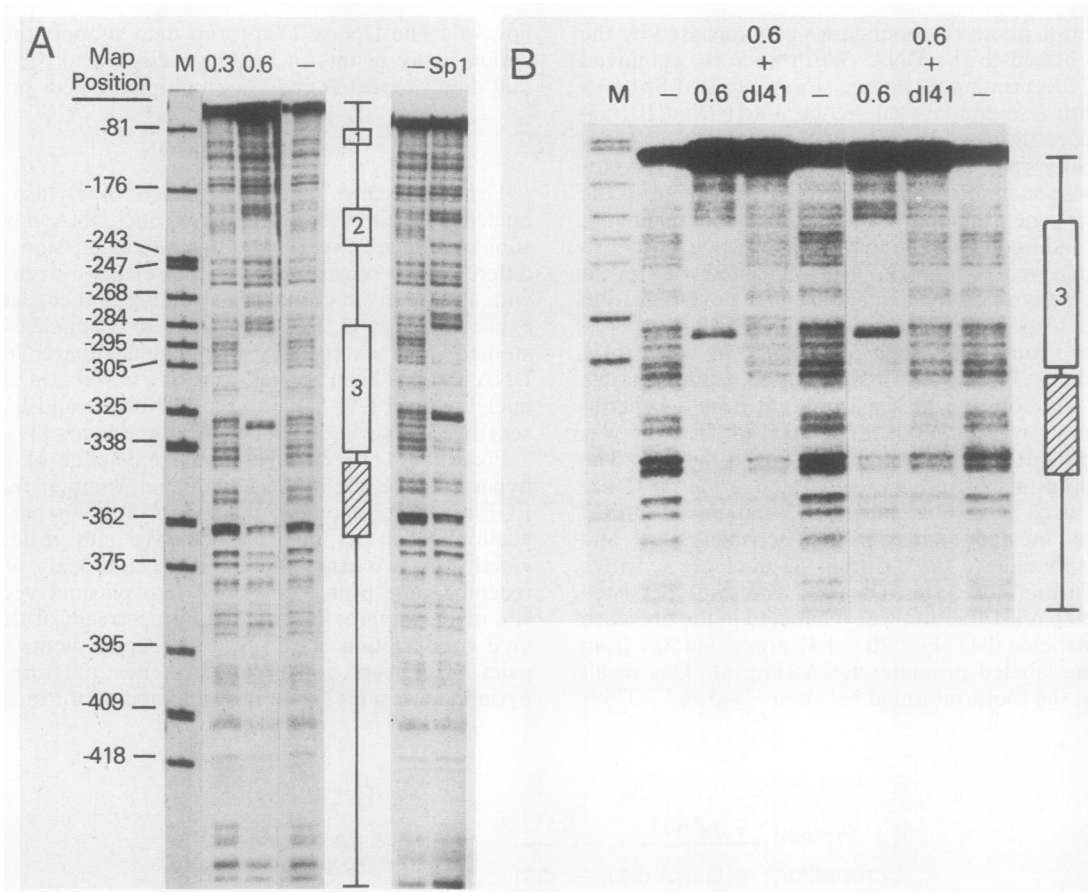


FIG. 9. DNase I footprint analysis of EGF receptor gene promoter (A) A431 nuclear extract was chromatographed over heparin-agarose, and NaCl elution fractions of 0.3 and 0.6 M (lanes 0.3 and 0.6) were collected. Protein, either A431 fractions (lanes 0.3 and 0.6) or purified Sp1 (lane Sp1), was incubated with a 469-bp *Mst*II-*Hin*dIII promoter fragment that was 5′ end labeled at the *Mst*II site (at position −485) and digested with DNase I. DNA fragments were resolved by denaturing polyacrylamide gel electrophoresis. Relative map positions of marker DNA fragments are shown at the left. Lane M, *Msp*I-digested pBR322 marker DNA (sizes from 67 to 404 bp); lanes −, no protein added. Boxes 1 to 3 represent known Sp1-binding sites (24). Box 1 is the weakest Sp1-binding site (24) and is not readily visible in this figure. The hatched box indicates the location of an additional 0.6 M NaCl-specific DNA footprint corresponding to direct repeat A. Strong DNase I-hypersensitive sites can be seen at positions −330 and −280 (lanes 0.6 and Sp1). Autoradiograms on the left (first four lanes) and right (last two lanes) were from two separate experiments. (B) Visualization of TCF DNase I footprint by prevention of Sp1 binding by direct competition. Extracts and probe were prepared as described for panel A. Two different preparations of the A431 0.6 M-NaCl heparin-agarose fraction were used. Lane M, *Msp*I-digested pBR322 marker DNA; lanes −, no protein added; lane 0.6, 0.6 M-NaCl fraction added; lanes 0.6 + dl41, 0.6 M-NaCl fraction and a 30-fold molar excess of the Sp1-binding 180-bp dl41 fragment (39) added.

general motif: TCCTCCTCC. This pyrimidine stretch is reminiscent of S1 nuclease-sensitive sites found in eucaryotic and viral DNAs, including promoter regions (18, 20, 34, 35, 55), and in the synthetic sequence $(dC\text{-}dT)_n \cdot (dA\text{-}dG)_n$ cloned into plasmid DNA (42).

We believe that the presence of direct repeats in the vicinity of the S1 nuclease-sensitive site is not a coincidence. It had been postulated that non-B DNA structures can exist when a DNA sequence contains either inverted repeats, capable of forming cruciform structures (32, 41), or direct repeats, capable of forming slipped structures (18, 34, 35, 55). We propose that the direct repeats in the EGF receptor gene promoter are sensitive to nuclease S1 because of the sliding of DNA strands past each other, after which a slipped structure is stabilized by base pairing between the complementary copies of the direct repeat. The only proof for the existence of such a structure in the EGF receptor gene is the generation of S1 nuclease promoter mutations missing one of the putative loops (Fig. 1). More rigorous proof for the formation of slipped structures has been presented for the adenovirus late promoter (55). Another possibility is that this homopyrimidine-homopurine stretch of DNA forms a triple-helical or triplex structure (28, 49).

S1 nuclease-sensitive sites associated with changes in promoter DNA secondary and tertiary structures have been implicated in differential gene expression (12, 29, 34, 40, 44, 55). We report that the direct-repeat elements in the EGF receptor gene promoter appear to be necessary for optimal transcription activity. Removal of the direct repeats by *Bal* 31 or S1 nuclease reduced CAT activity in a transient-transfection assay from two- to fivefold (Fig. 3 and 4). Depression in CAT activity resulting from deletion of an S1-sensitive site has also been observed in the α2(I) collagen gene promoter (C. McKeon, personal communication).

The exact mechanism by which nuclease-sensitive sites are involved in regulating gene expression is unclear. One hypothesis is that their single-stranded nature allows RNA polymerase entry and the resulting formation of an open transcription complex (50).

Another exciting possibility is that local changes in DNA

structure, as detected by DNase I and S1 nuclease digestion, alter the interactions between promoter DNA and specific DNA-binding protein. It was, therefore, of great interest to determine whether the S1 nuclease-sensitive site in the EGF receptor gene promoter is associated with a DNA-binding protein. Both gel retardation and DNase I footprinting assays were performed by using labeled EGF receptor gene promoter fragments and nuclear extracts from A431 cells. These experiments clearly showed that two proteins were able to bind to repeat A, which was responsible for conferring S1 nuclease sensitivity to the promoter (Fig. 6 to 9). One protein is the positive transcription factor Sp1 (2, 8, 9, 25). The other protein is what we refer to as TCF. Upon obliteration of the structure of repeat A, neither Sp1 nor TCF bound to the region (Fig. 7). This result establishes a correlation between the binding of these proteins and optimal transcription activity as determined by transient-transfection assays. We believe that this paper is the first published report correlating S1 nuclease sensitivity, *trans*-acting-factor binding, and altered gene expression.

The EGF receptor gene promoter S1 nuclease-sensitive site consisted of the two pairs of direct repeats A and B (Fig. 1); all four conformed closely to the motif TCCTCCTCC, and all four contained potential Sp1-binding sites. The relationship between the two pairs of repeats may be significant. An unlabeled DNA fragment containing both sets of direct repeats competed for binding about ten times more effectively in the gel retardation assay than did a competitor DNA fragment possessing only the upstream set of direct repeats (repeat A) (Fig. 6 and 7). The results of these competition experiments raised the possibility that the same or similar factors bind to both sets of repeats. This hypothesis is likely the case for Sp1, which bound strongly to the downstream set of repeats (repeat B) (Fig. 9). Furthermore, it is conceivable that Sp1 and TCF interact, perhaps in a dynamic or cooperative fashion. This possibility is most intriguing because the four repeat elements are separated from each other by 19, 21, and 23 bp, or by about two full turns of the DNA helix. The exact relationship between the binding of these factors and the structure of the DNA in this region is presently unknown.

Sequences resembling the human EGF receptor gene promoter TCCTCCTCC region are found in several other human receptor gene promoters, including the insulin and low-density lipoprotein (LDL) receptor genes (1, 46). Although S1 nuclease sensitivity has not been examined in these promoters, the TC-rich region in the LDL receptor gene is part of a sterol regulatory element and binds a nuclear protein (46). It remains to be established whether other S1 nuclease-sensitive promoter regions containing DNA sequences similar to that of the EGF receptor gene promoter TCCTCCTCC motif are associated with the binding of Sp1, TCF, or related *trans*-acting factors.

ACKNOWLEDGMENTS

We wish to thank Ira Pastan, Ryoichiro Kageyama, Pamela Marino, and Catherine McKeon for useful and stimulating discussions. We are indebted to Cliff Parkison for dideoxy sequencing, James Kadonaga and Robert Tjian for purified Sp1, John Brady for pdl41, Bruce Paterson for pβA-CAT, and Ryoichiro Kageyama for A431 nuclear extract chromatographic fractions BA and BC. We acknowledge Steven Neal for photographic assistance and Jennie Evans and Althea Gaddis for editorial assistance.

LITERATURE CITED

1. **Araki, E., F. Shimada, M. Uzawa, M. Mori, and Y. Ebina.** 1987. Characterization of the promoter region of the human insulin receptor gene. Evidence for promoter activity. J. Biol. Chem. **262**:16186–16191.
2. **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.
3. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
4. **Cohen, S., H. Ushiro, C. Stoscheck, and M. Chinkers.** 1982. A native 170,000 epidermal growth factor receptor-kinase complex from shed plasma membrane vesicles. J. Biol. Chem. **257**:1523–1531.
5. **Covelli, I., R. Mozzi, R. Rossi, and L. Frati.** 1972. The mechanism of action of the epidermal growth factor III. Stimulation of the uptake of labeled precursors into RNA, DNA and proteins induced by EGF in isolated tumor cells. Hormones **3**:183–191.
6. **Cowley, G., J. A. Smith, B. Gusterson, F. Hendler, and B. Ozanne.** 1984. The amount of EGF receptor is elevated on squamous cell carcinomas. Cancer Cells **1**:5–10.
7. **Downward, J., Y. Yarden, E. Mayes, G. Scarce, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger, and M. D. Waterfield.** 1984. Close similarity of epidermal growth factor receptor and v-erbB oncogene protein sequence. Nature (London) **307**:521–527.
8. **Dynan, W. S., and R. Tjian.** 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell **32**:669–680.
9. **Dynan, W. S., and R. Tjian.** 1983. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell **35**:79–87.
10. **Elgin, S. C. R.** 1981. DNase I-hypersensitive sites of chromatin. Cell **27**:413–415.
11. **Elgin, S. C. R.** 1982. Chromatin structure, DNA structure. Nature (London) **300**:402–403.
12. **Evans, T., E. Schon, G. Gora-Maslak, J. Patterson, and A. Efstradiatis.** 1984. S1-hypersensitive sites in eukaryotic promoter regions. Nucleic Acids Res. **12**:8043–8058.
13. **Fabricant, R. N., J. E. Delarco, and G. J. Todaro.** 1977. Nerve growth factor receptor on human melanoma cells in culture. Proc. Natl. Acad. Sci. USA **74**:565–569.
14. **Goding, C. R., and W. C. Russell.** 1983. S1 sensitive sites in adenovirus DNA. Nucleic Acids Res. **11**:21–36.
15. **Gorman, C., L. Moffat, and B. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
16. **Gorman, C. M., 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.
17. **Haley, J., N. Whittle, P. Bennett, D. Kinchington, A. Ullrich, and M. Waterfield.** 1987. The human EGF receptor gene: structure of the 110 kb locus and identification of sequences regulating its transcription. Oncogene Res. **1**:375–396.
18. **Hentschel, C. C.** 1982. Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to the S1 nuclease. Nature (London) **295**:714–716.
19. **Hollenberg, M. D., and P. Cuatrecasas.** 1973. Epidermal growth factor: receptors in human fibroblasts and modulation of action by cholera toxin. Proc. Natl. Acad. Sci. USA **70**:2964–2968.
20. **Htun, M., E. Lund, and J. E. Dahlberg.** 1984. Human U1 RNA genes contain an unusual nuclease S1 cleavage site within the conserved 3′ flanking region. Proc. Natl. Acad. Sci. USA **81**:7288–7292.
21. **Ishii, S., J. T. Kadonaga, R. Tjian, J. N. Brady, G. T. Merlino, and I. Pastan.** 1986. Binding of the Sp1 transcription factor by the human Harvey ras 1 proto-oncogene promoter. Science **232**:1410–1413.

22. **Ishii, S., G. T. Merlino, and I. Pastan.** 1985. Promoter region of the human Harvey ras proto-oncogene: similarity to the EGF receptor proto-oncogene promoter. Science **230**:1378–1381.
23. **Ishii, S., Y.-H. Xu, R. M. Stratton, B. A. Roe, G. T. Merlino, and I. Pastan.** 1985. Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proc. Natl. Acad. Sci. USA **82**:4920–4924.
24. **Johnson, A. C., S. Ishii, Y. Jinno, I. Pastan, and G. T. Merlino.** 1988. EGF receptor gene promoter: deletion analysis and identification of nuclear protein binding sites. J. Biol. Chem. **263**: 5693–5699.
25. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83**:5889–5893.
26. **Kageyama, R., G. T. Merlino, and I. Pastan.** 1988. Epidermal growth factor (EGF) receptor gene transcription: requirement for Sp1 and an EGF receptor-specific factor. J. Biol. Chem. **263**: 6329–6336.
27. **King, C. R., M. H. Kraus, L. T. Williams, G. T. Merlino, I. H. Pastan, and S. A. Aaronson.** 1985. Human tumor cell lines with EGF receptor gene amplification in the absence of aberrant sized mRNAs. Nucleic Acid Res. **13**:8477–8486.
28. **Kohwi-Shigematsu, T., N. Scribner, and Y. Kohwi.** 1988. An ultimate chemical carcinogen, N-acetoxy-2-acetylaminofluorene detects non-B DNA structures that are reactive with chloroacetaldehyde in supercoiled plasmid DNA. Carcinogenesis **9**:457–461.
29. **Larsen, A., and M. Weintraub.** 1982. An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell **29**:609–622.
30. **Lembach, K. J.** 1976. Enhanced synthesis and extracellular accumulation of hyaluronic acid during stimulation of quiescent human fibroblasts by mouse epidermal growth factor. J. Cell. Physiol. **89**:277–288.
31. **Libermann, T. A., H. R. Nasbaum, N. Razon, R. Kris, I. Lax, H. Soreg, N. Whittle, M. D. Waterfield, A. Ullrich, and J. Schlessinger.** 1985. Amplification, enhanced expression and possible rearrangment of EGF receptor gene in primary human brain tumors of glial origin. Nature (London) **313**:144–147.
32. **Lilley, D. M. F.** 1980. The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc. Natl. Acad. Sci. USA **77**:6468–6472.
33. **Lin, C. R., W. S. Chen, W. Kruiger, L. S. Stolarsky, W. Weber, R. M. Evans, I. M. Verma, G. N. Gill, and M. G. Rosenfeld.** 1984. Expression cloning of human EGF receptor complementary DNA: gene amplification and three related messenger RNA products in A431 cells. Science **224**:843–848.
34. **Mace, H. A. F., H. R. B. Pelham, and A. A. Travers.** 1983. Association of an S1 nuclease sensitive structure with short direct repeats 5′ of Drosophila heat shock genes. Nature (London) **304**:555–557.
35. **McKeon, C., A. Schmidt, and B. deCrombrugghe.** 1984. A sequence conserved in both the chicken and the mouse α2(I) collagen promoter contains sites sensitive to S1 nuclease. J. Biol. Chem. **259**:6636–6640.
36. **Merlino, G. T., J. S. Tyagi, B. deCrombrugghe, and I. Pastan.** 1982. Transcription of the chicken α2 (type I) collagen gene by homologous cell-free extracts. J. Biol. Chem. **257**:7254–7261.
37. **Merlino, G. T., Y.-H. Xu, S. Ishii, A. J. L. Clark, K. Semba, K. Toyoshima, T. Yamamoto, and I. Pastan.** 1984. Amplification and enhanced expression of the epidermal growth factor receptor gene in A431 human carcinoma cells. Science **224**:417–419.
38. **Merlino, G. T., Y.-H. Xu, N. Richert, A. J. L. Clark, S. Ishii, S. Banks-Schlegal, and I. Pastan.** 1985. Elevated epidermal growth factor receptor gene copy number and expression in a squamous carcinoma cell line. J. Clin. Invest. **75**:1077–1079.
39. **Mishoe, M., J. N. Brady, M. Radonovich, and N. P. Salzman.** 1984. Simian virus 40 guanine-cytosine-rich sequences function as independent transcriptional control elements in vitro. Mol. Cell. Biol. **4**:2911–2920.
40. **Nickol, J. M., and G. Felsenfeld.** 1983. DNA conformation at the 5′ end of the chicken adult β-globin gene. Cell **35**:467–477.
41. **Panayotatos, N., and R. D. Wells.** 1981. Cruciform structures in supercoiled DNA. Nature (London) **289**:466–470.
42. **Pulleyblank, D. E., D. B. Haniford, and A. R. Morgan.** 1985. A structural basis for S1 nuclease sensitivity of double-stranded DNA. Cell **42**:271–280.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. RNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
44. **Schon, E., T. Evans, J. Welsch, and A. Efstratiadis.** 1983. Conformation of a promoter DNA: fine structure mapping of S1-hypersensitive sites. Cell **35**:837–848.
45. **Singh, H., S. Ranjan, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
46. **Sudhof, T. C., D. W. Russell, M. S. Brown, and J. L. Goldstein.** 1987. 42 bp element from LDL receptor gene confers end-product repression by sterols when inserted into viral TK promoter. Cell **48**:1061–1069.
47. **Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. M. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg.** 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature (London) **309**:418–425.
48. **Velu, T. J., L. Beguinot, W. C. Vass, M. C. Willingham, G. T. Merlino, I. Pastan, and D. R. Lowy.** 1987. Epidermal growth factor dependent transformation by a human EGF receptor proto-oncogene. Science **238**:1408–1410.
49. **Voloshin, O. N., S. M. Mirkin, V. I. Lyamichev, B. P. Belotserkovskii, and M. D. Frank-Kamenetskii.** 1988. Chemical probing of homopurine-homopyrimidine mirror repeats in supercoiled DNA. Nature (London) **333**:475–476.
50. **Weintraub, H.** 1983. A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. Cell **32**:1191–1203.
51. **Weintraub, H.** 1985. High-resolution mapping of S1- and DNase I-hypersensitive sites in chromatin. Mol. Cell. Biol. **5**:1538–1539.
52. **Weisbrod, S.** 1982. Active chromatin. Nature (London) **297**:289–295.
53. **Xu, Y.-H., S. Ishii, A. J. L. Clark, M. Sullivan, R. K. Wilson, D. P. Ma, B. A. Roe, G. T. Merlino, and I. Pastan.** 1984. Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs overproduced in A431 carcinoma cells. Nature (London) **309**:806–810.
54. **Xu, Y.-H., N. Richert, S. Ito, G. T. Merlino, and I. Pastan.** 1984. Characterization of epidermal growth factor receptor gene expression in malignant and normal human cell lines. Proc. Natl. Acad. Sci. USA **81**:7308–7312.
55. **Yu, Y.-T., and J. L. Manley.** 1986. Structure and function of the S1 nuclease-sensitive site in the adenovirus late promoter. Cell **45**:743–751.