Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 4920–4924, August 1985
Biochemistry

Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene

(transcription/oncogene/DNase I hypersensitivity/simian virus 40/hydroxymethylglutaryl-CoA reductase)

SHUNSUKE ISHII*, YOUNG-HUA XU*, RANDY H. STRATTON†, BRUCE A. ROE†, GLENN T. MERLINO*, AND IRA PASTAN*‡

*Laboratory of Molecular Biology, Division of Cancer, Biology and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205; and †Department of Chemistry, University of Oklahoma, Norman, OK 73019

Contributed by Ira Pastan, April 4, 1985

ABSTRACT **The promoter region of the epidermal growth factor (EGF) receptor has been identified by *in vitro* transcription using EGF receptor genomic DNA fragments as template and by primer extension and nuclease S1 mapping using EGF receptor mRNA. Six transcriptional start sites were identified. DNA sequence analysis shows that the promoter region contains neither a "TATA box" nor a "CAAT box," has an extremely high G+C content (88%), and contains five CCGCCC repeats and four (TCC)TCCTCCTCC repeats. This promoter region is situated close to or within a DNase I-hypersensitive site in A431 human epidermoid carcinoma cells, which overproduce the EGF receptor. The EGF receptor gene promoter has some resemblance to the promoter of the hydroxymethylglutaryl-CoA reductase gene and the early promoter of simian virus 40. This similarity may offer a clue to the mechanism by which the receptor gene is regulated.**

Epidermal growth factor (EGF) stimulates cell growth by binding to specific membrane receptors (1, 2). The EGF receptor is a membrane-spanning glycoprotein that has three functional domains: an EGF-binding domain located on the external cell surface, a transmembrane domain, and a cytoplasmic tyrosine kinase domain (3–5). EGF binding induces internalization of the receptor, which is ultimately delivered to lysosomes where it is destroyed (6), and induces accumulations of EGF receptor mRNA (unpublished observations). The binding of EGF to the receptor activates a tyrosine kinase that phosphorylates various cellular proteins, including the EGF receptor itself (7, 8). The best characterized tyrosine-specific protein kinase is $pp60^{src}$, the transforming protein of the Rous sarcoma virus (9–11). The oncogene *erbB*, carried by the avian erythroblastosis retrovirus (AEV), is homologous to the transmembrane and kinase domains of the EGF receptor gene (12–17). The EGF receptor is often overproduced in tumor cells (18); the human epidermoid carcinoma A431, a squamous carcinoma cell line, and several glioblastomas contain a very large number of EGF receptors resulting from EGF receptor gene amplification (14, 15, 17, 19–21). Therefore, it seems plausible that overexpression of the EGF receptor gene contributes to the phenotype of cellular transformation. To investigate the regulation of the expression of the EGF receptor gene at the molecular level, we identified and characterized the 5′ flanking region of this gene.

MATERIALS AND METHODS

DNA analysis and sequencing were performed as described (22–25). Poly(A)$^+$ RNA was isolated from A431 and epithelial cells (26). Extension of various primers was performed as described (27, 28). Nuclease S1 mapping was done according to a modification of procedures detailed elsewhere (28–30). For *in vitro* transcription studies, HeLa cell extracts were prepared and incubations were performed as previously reported (27, 31). DNase I-hypersensitive sites were detected as described (32); the nuclei were isolated according to Wu *et al.* (33), including centrifugation through a 1.5 M sucrose cushion at 100,000 × *g* for 30 min.

RESULTS

Isolation of 5′-Specific Human EGF Receptor Genomic Clones. Genomic clones containing the 5′ end of the EGF receptor gene were isolated from a human genomic DNA fetal liver library (34) by screening with a mixture of synthetic oligonucleotide probes [CTGGA(A/G)GA(A/G)AA(A/G)AA] that contains the possible codons of the amino acids at the NH_2 terminus of the EGF receptor (Leu-Glu-Glu-Lys-Lys-) (35). One positive clone was isolated (λER1). A 950-base-pair (bp) *Eco*RI–*Sst* I (nucleotides −965 to −15) fragment of λER1 was found to hybridize with a 140-bp *Cla* I–*Sst* I (nucleotides −155 to −15) DNA fragment from the cDNA clone pE15 (36) that contains the 5′ untranslated sequences of the EGF receptor gene (Fig. 1). Additional genomic clones were selected in a second screening with this 950-bp *Eco*RI–*Sst* I fragment as a probe. Two other clones, λER9 and λER14, were isolated and characterized (Fig. 1).

Sequence analysis of both the 950-bp *Eco*RI–*Sst* I fragment and the adjacent 3′ region revealed that one exon is located in this area; this exon contains the 5′ untranslated region, 72 bp coding for the signal sequence, and 16 bp coding for the NH_2-terminal portion of the mature EGF receptor (Fig. 1). The exon ends with the sequence AG and the adjacent intron begins with the sequence GT.

Primer Extension. To identify the 5′ boundary of the EGF receptor RNA, a ^{32}P-labeled *Ava* I–*Hpa* II (Fig. 1, nucleotides −101 to −25) single-stranded DNA primer was prepared from plasmid pE15 and hybridized with A431 poly(A)$^+$ RNA. The primer was extended with avian myeloblastosis virus reverse transcriptase, and the sizes of the resulting products were determined by gel electrophoresis under denaturing conditions, followed by autoradiography. The most abundant extension products had lengths of 234, 160, 141, 93, 87, and 83 bases (Fig. 2*a*, bands 1–6, respectively). These results suggest that the start site of EGF receptor gene transcription may be heterogeneous, occurring 258, 184, 165, 117, 111, and 107 bp upstream from the initiator AUG codon (Fig. 1).

One possibility is that many of these sites are unique to A431 cells, which possess amplified and rearranged EGF

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EGF, epidermal growth factor; HMG-CoA reductase, hydroxymethylglutaryl-CoA reductase; SV40, simian virus 40; bp, base pair(s); kb, kilobase(s).

‡To whom reprint requests should be addressed at: Building 37, Rm. 4E-16, National Institutes of Health, Bethesda, MD 20205.

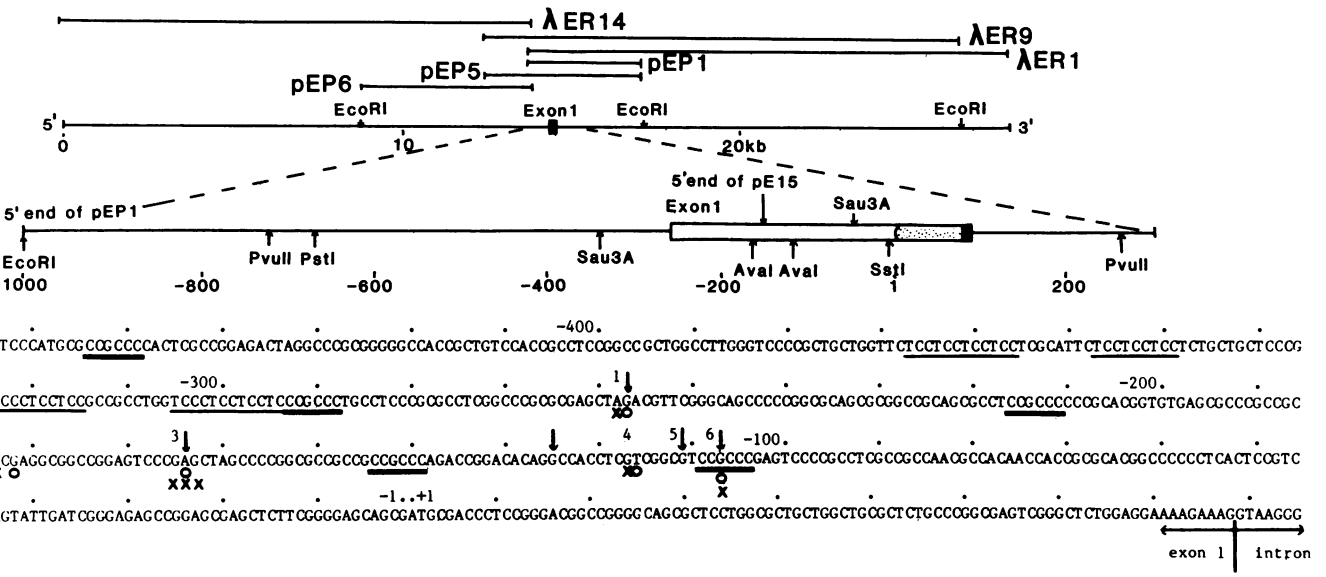

FIG. 1. Structure of the human EGF receptor genomic clones. (*Upper*) A schematic representation of the three overlapping EGF receptor bacteriophage λ recombinants is at the top. The subcloned genomic fragments are shown below the λ recombinants (pEP1, pEP5, and pEP6). Below the *Eco*RI restriction map is an expanded map of the promoter and exon 1 region. Untranslated sequences are represented by a line, the exon by a box. The unshaded area of the box represents sequences that encode the 5′ untranslated region; the stippled area of the box codes for the signal peptide; the shaded area of the box codes for the mature EGF receptor. (*Lower*) Nucleotide sequence of the 5′ portion of the human EGF receptor gene, where +1 corresponds to the A of the initiator methionine codon, and residues preceding it are represented by negative numbers. The multiple initiation sites (numbered 1–6) demonstrated by primer extension, nuclease S1 mapping, and *in vitro* transcription are indicated by x, o, and ↓, respectively. The repeated CCGCCC sequences are underlined by a thick solid line, whereas the repeated TCCTCCTCC sequences are underlined by a thin solid line.

receptor genes (14, 15, 17, 36). Normal human epithelial cell (37) poly(A)$^+$ RNA was therefore isolated and primer-extended. The same population of extended products is synthesized whether template RNA is from normal cells (Fig. 2*c*, lane D) or A431 cells (lane F). Therefore, the 5′ end of the EGF receptor gene of A431 cells resembles that of a normal cell.

Nuclease S1 Mapping. When a 288-bp *Sau*3A DNA fragment from genomic subclone pEP1 (Fig. 1, nucleotides −324 to −36) was 5′-end-labeled, hybridized with A431 poly(A)$^+$ RNA, and digested with nuclease S1, multiple protected fragments were detected (Fig. 2*b*). The major fragments had lengths of 222, 148, 130, and 82 bases (Fig. 2*b*, bands 1–4), corresponding to initiation sites at nucleotides −257, −182, −164, and −116. Two fragments corresponding to the two distal start sites were very weak but were seen after longer autoradiographic exposure.

DNA Sequence Comparison. To determine whether the 5′ ends of the extended primers originate from within λER1 genomic DNA, we compared the nucleotide sequence of an extended DNA primer to that of the cloned genomic DNA. By using a labeled synthetic oligonucleotide primer (30-mer) complementary to one part of the cDNA sequence (Fig. 2*c*, −176 to −147) a single major labeled extension product was synthesized (Fig. 2*c*, lane E). The nucleotide sequence of this extended primer was found to be identical to the sequence of the genomic clone immediately adjacent to the primer, except for position −214 (Fig. 2*d*, asterisk). The EGF receptor cDNA of A431 cells has a T; the genomic clone from fetal liver has a G. These results indicate that the genomic DNA corresponding to the extended cDNA primer is not interrupted by introns.

The primer extension, nuclease S1 mapping, and sequencing data together strongly suggest that the synthesized cDNAs extend to the extreme 5′ end of the EGF receptor gene. Ullrich *et al.* (14) isolated an EGF receptor cDNA clone from an A431 library which contains 176 bp of 5′ untranslated cDNA. However, the 5′-most 10-bp sequence of their cDNA clone (5′ GCCGCGCTGC 3′) is not found in the sequence of either the extended cDNA primer made from A431 RNA or the λER1 genomic clone isolated from human fetal liver. We did find an inverted sequence (5′ GCAGCGCGGC 3′) of this 10-bp region at nucleotide position −236 to −227.

Multiple EGF Receptor mRNAs Initiate from the Same 5′-Most Site. A431 cells contain an aberrant 2.8- to 2.9-kilobase (kb) EGF receptor RNA which diverges from EGF receptor RNAs found in other cell types (10 and 5.6 kb) at its 3′ end (14, 16, 17, 36). To determine whether different A431 EGF receptor RNA species are transcribed from variant start sites, we performed RNA blot analysis by using a 641-bp *Pst* I–*Ava* I probe (Fig. 1, −626 to −165) specific for the 5′ portion of the 5′ untranslated region; this DNA can only hybridize to transcripts starting from the two initation sites at nucleotide positions −258 and −184 (Fig. 1). The 461-bp probe hybridized to all of the A431 EGF receptor RNA species (10, 5.6, and 2.9 kb; Fig. 3, lane a). A *Cla* I–*Sst* I 140-bp DNA fragment (Fig. 1, −159 to −19) of the pE15 cDNA clone, containing the 3′ portion of the 5′ untranslated region, also hybridized with all of the RNA species (Fig. 3, lane b). These results show that all three RNAs are at least transcribed from the start sites at nucleotides −258 and −184 and subsequently processed into the three different-size RNA species.

Sequence Around the EGF Receptor Promoter Is G+C-Rich. The nucleotide sequence upstream of the EGF receptor RNA start sites contains neither a "TATA box" nor a "CAAT box" (Fig. 1). Most other characterized eukaryotic genes have these two sequence elements ≈30 and 80 bp, respectively, upstream of the RNA initiation site (38). Since *in vitro* transcriptional studies have shown that the TATA box serves to fix the site at which transcription will start (39), the fact that the EGF receptor gene has numerous transcriptional initiation sites is not surprising.

The sequence between nucleotides −540 and −1 (Fig. 1), which contains the putative promoter and 5′ untranslated region, has a G+C content of 88%; however, the region further upstream of −540 has a G+C content of only about 50%. Further analysis of the 5′ flanking region of the EGF

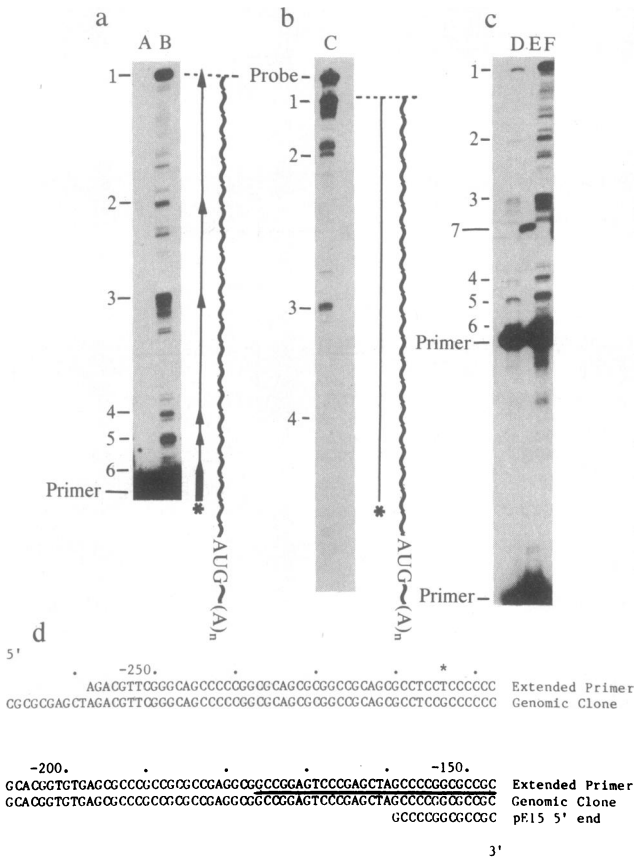

FIG. 2. Determination of the EGF receptor gene transcription start sites. (*a*) Primer extension of the A431 EGF receptor RNAs. Five micrograms of A431 poly(A)$^+$ RNA (lane B) or 5 μg of calf liver tRNA (lane A) was hybridized with a 5′ end-labeled (*) single-stranded *Ava* I–*Hpa* II (nucleotides −101 to −25) DNA fragment (thick line at right). Primer-extended products (1–6, arrows) were analyzed in a 7 M urea/8% polyacrylamide gel. RNA is represented by a wavy line containing the downstream translation start site (AUG) and a poly(A) tail. (*b*) Nuclease S1 mapping of the 5′ portion of EGF receptor RNAs. A431 poly(A)$^+$ RNA (5 μg; wavy line) was hybridized to an end-labeled (*) *Sau*3A 288-bp (nucleotides −324 to −36) DNA probe, and the nuclease S1-resistant products (1–4, solid line) were electrophoresed. The appearance of the relatively strong band above fragment 2 is caused by nuclease S1 cleavage of the mismatched base pair generated by hybridizing A431 RNA to fetal liver genomic DNA (see Fig. 2*d*) and is not indicative of a transcriptional start site. (*c*) Primer extension of the human EGF receptor RNAs. Fifty micrograms of epithelial cell (lane D) or 5 μg of A431 cell poly(A)$^+$ RNA (lanes E and F) was hybridized with either the *Ava* I–*Hpa* II pEP1 primer as in *a* (lanes D and F) or with a 5′-end-labeled synthetic 30-mer oligonucleotide complementary to nucleotides −176 to −147 of the genomic sequence (lane E). (*d*) DNA sequence of the extended primer (band 7 in lane E). The asterisk represents the single base difference between the extended primer and the genomic clones. The original 30-mer primer is underlined.

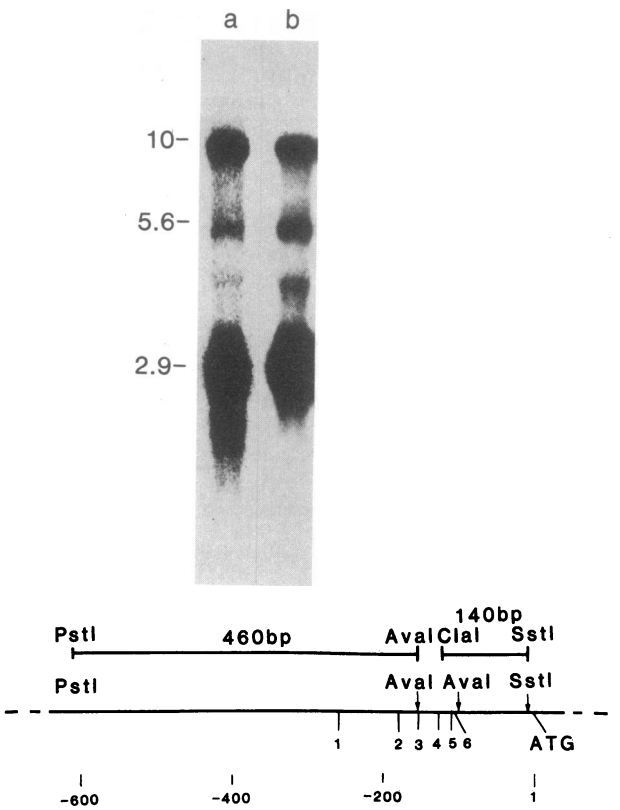

FIG. 3. RNA blot analysis of EGF receptor RNA. A431 cell poly(A)$^+$ RNA was fractionated in 1% agarose containing formaldehyde and then was transferred to nitrocellulose. Probes were the 461-bp *Pst* I–*Ava* I DNA fragment (nucleotides −626 to −165), which covers about half of the 5′ untranslated region (lane a) and the *Cla* I–*Sst* I 140-bp pE15 DNA fragment (nucleotides −159 to −19) (lane b). Numbers at left indicate the size (in kb) of the major RNA species. The restriction enzyme map of the 5′ untranslated region is shown at the bottom. The transcriptional start sites 1–6 are indicated. The two hybridization probes are depicted above the genomic map.

receptor gene shows that a specific 6-bp sequence (CCGCCC) is repeated five times in the putative promoter region at −453, −293, −216, −144, and −109 (Fig. 1). CCGCCC is the same sequence that is repeated six times within the simian virus 40 (SV40) early promoter (39). Another repeated sequence [(TCC)TCCTCCTCC] was found at −367, −347, −324, and −305. Similar sequences have been found in the upstream promoter region of both chicken and mouse α2(I) collagen genes and occur at sites sensitive to nuclease S1 (40).

***In Vitro* Transcription of the EGF Receptor Gene.** To determine whether the cloned genomic DNA containing the 5′ flanking region of the EGF receptor gene is active in initiating transcription, we performed an *in vitro* transcription assay using extracts from HeLa cells (31). The plasmid pEP1 was truncated, with *Sst* I, downstream from the initiation start sites and then used as template (Fig. 4*a*). *Sst* I cuts at a site 20 bp upstream of the EGF receptor protein initiation codon AUG (Fig. 1). Transcription products were fractionated in a 7 M urea/8% polyacrylamide gel. Three major (235, 165, and 144 bases) and three minor (105, 91, and 87 bases) α-amanitin-sensitive transcripts were observed (Fig. 4*b*). Thus RNA was initiated 257, 184, 164, 125, 111, and 107 bp upstream of the translation start site, corroborating *in vivo* results previously obtained by primer extension and nuclease S1 mapping (Fig. 4*c*).

DNase I-Hypersensitive Sites. Transcriptionally active chromatin contains discrete DNA regions that are hypersensitive to DNase I and are often located near the 5′ portion of genes (41). If the EGF receptor promoter has been accurately localized, this region should be very sensitive to DNase I in cells actively synthesizing the EGF receptor. We determined the DNase I-hypersensitive sites around the putative promoter region in A431 cells, which overproduce EGF receptors (19), and in HTB54 cells, which do not produce the EGF receptor.

Nuclei were isolated and exposed to various amounts of DNase I. The genomic DNA was then isolated, digested with *Eco*RI, fractionated electrophoretically in 1% agarose, and transferred to nitrocellulose. Indirect "end-labeling" was utilized to visualize hypersensitive sites at the 5′ end of the EGF receptor gene. A 400-bp *Pvu* II–*Eco*RI DNA fragment corresponding to the extreme 3′ end of the promoter-bearing

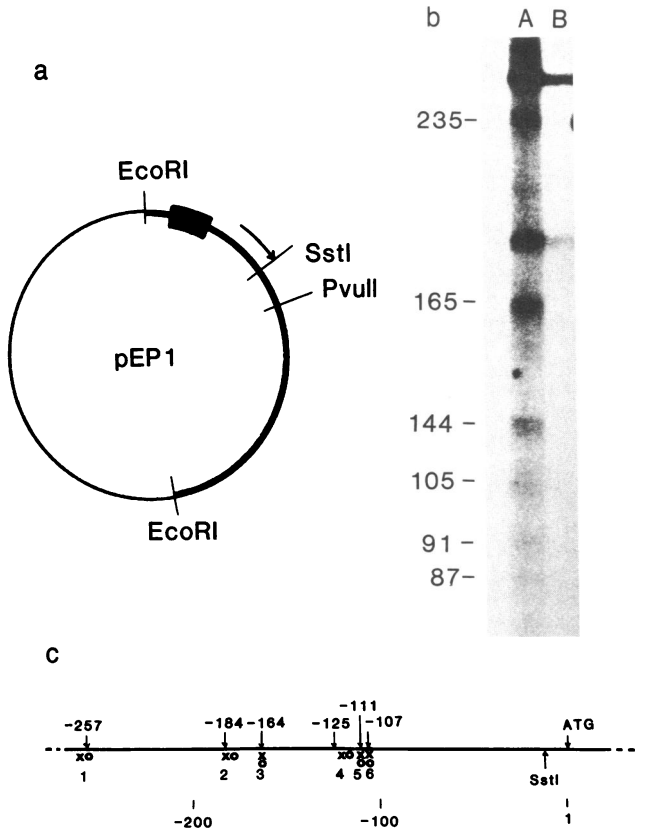

FIG. 4. *In vitro* transcription of the EGF receptor gene. (*a*) The pEP1 plasmid linearized with *Sst I* was used as a template for *in vitro* transcription. The shaded box represents the promoter region; the arrow indicates the direction of transcription. (*b*) RNAs were synthesized using [α-^{32}P]GTP and cellular extracts with (lane B) or without (lane A) α-amanitin (1 μg/ml) and then were fractionated in a 7 M urea/8% polyacrylamide gel. Numbers at left indicate the size (in nucleotides) of the α-amatinin-sensitive RNA polymerase II transcripts. (*c*) Summary of the location of transcriptional initiation sites by primer extension (x), nuclease S1 mapping (○), and *in vitro* transcription (arrows).

*Eco*RI 8-kbp genomic fragment (Fig. 5) was used to map the hypersensitive sites. In addition to the *Eco*RI 8-kb fragment, this probe hybridized to a 4-kbp and a 1.8-kbp A431 genomic DNA fragment (Fig. 5*a*). This indicates that one hypersensitive site is localized 100–300 bp upstream from the most 5′ major transcriptional start site and that the other hypersensitive site is within the first intron (Fig. 5). Similar DNase I-hypersensitivity results have been obtained for chicken chromatin containing the α2(I) collagen gene (32, 42). When HTB54 nuclei were used, no obvious bands were generated by DNase I digestion (Fig. 5*b*).

To confirm the location of these hypersensitive sites, a probe that corresponds to the extreme 5′ end of the *Eco*RI 8-kb genomic fragment, an *Eco*RI–*Hin*dIII 2.2-kbp DNA fragment, was utilized. In addition to the 8-kbp fragment, this probe hybridized to 6.6- and 4.3-kbp fragments of A431 genomic DNA (Fig. 5*c*), but not to any fragments of HTB54 DNA (Fig. 5*d*). These data confirm the location of the two hypersensitive sites.

DISCUSSION

We have localized the EGF receptor gene "promoter" in cloned human genomic DNA. The EGF receptor gene does not possess typical TATA or CAAT boxes, and initiation of RNA transcription occurs at multiple sites. The 5′ flanking region of the EGF receptor gene therefore differs from the corresponding regions of most other eukaryotic genes analyzed to date but resembles the SV40 early promoter and the hydroxymethylglutaryl-CoA reductase (HMG-CoA reductase) promoter (Fig. 6). The SV40 early promoter has a G+C-rich region, about 80 bp upstream of the RNA start site, which is essential for transcription (39, 46–49) and contains six copies of the sequence CCGCCC (Fig. 6). The gene of HMG-CoA reductase does not possess a typical TATA box and has five transcriptional initiation sites (43). The 5′ flanking region of this gene is G+C-rich (65%) and has three repeats of the sequence CCGCCC (Fig. 6). The promoter region of the c-*myc* protooncogene also contains the CCGCCC sequence and is G+C-rich (66%; ref. 45) (Fig. 6). Like the corresponding region of each of these other genes, the 5′ flanking region of the EGF receptor gene is G+C-rich (88%). Within this area there are five repeats of the CCGCCC sequence; however, the sequence of the third repeat in the EGF receptor gene was found to vary between A431 and normal epithelial cells.

Recently, Gidoni *et al.* (50) reported that the human HeLa transcription factor Sp1 that stimulates the *in vitro* transcription of the SV40 early promoter binds to specific SV40 early gene regions containing CCGCCC sequences. Because Sp1 protein is found in uninfected cells, it seems likely that it plays a role in the expression of cellular as well as viral genes (51). An attractive hypothesis is that some "growth control" genes that contain the repeat CCGCCC in the promoter region are regulated by binding of the Sp1 protein. The

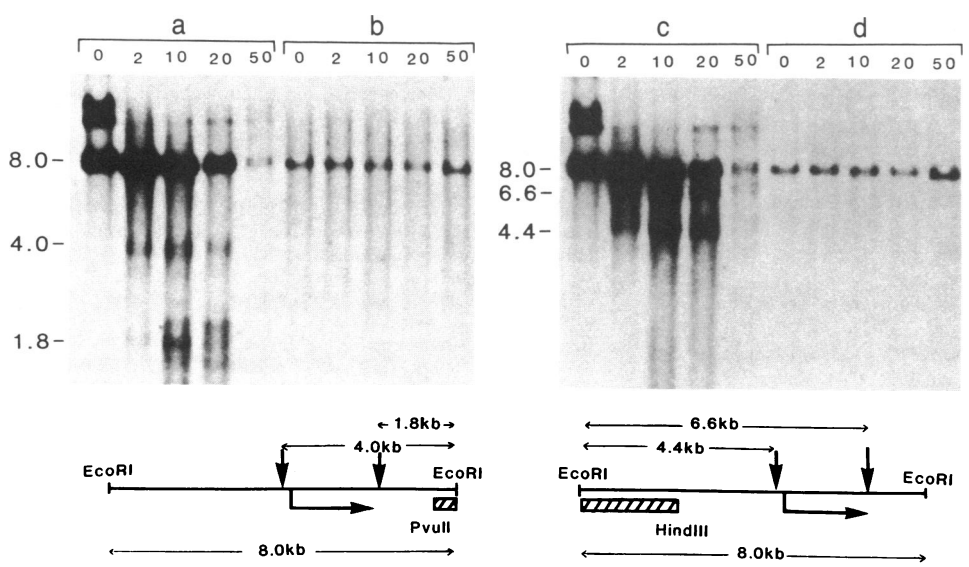

FIG. 5. Mapping of DNase I-hypersensitive sites near the 5′ portion of the EGF receptor gene. Nuclei of A431 (*a* and *c*) or HTB54 epidermoid carcinoma cells (*b* and *d*) were digested at 23°C for 5 min with DNase I at 0–50 units/ml (numbers at top). Two probes were utilized for hybridization: a *Pvu* II–*Eco*RI 400-bp DNA fragment (*a* and *b*) or an *Eco*RI–*Hin*dIII 2.2-kbp DNA fragment (*c* and *d*). Numbers at left indicate the size of the bands in kb. The positions of the DNase I-hypersensitive sites (large arrows) and the two nick-translated probes (hatched boxes) in the genomic clone of the EGF receptor gene are shown at the bottom. The bent arrows indicate the approximate start and the direction of transcription.

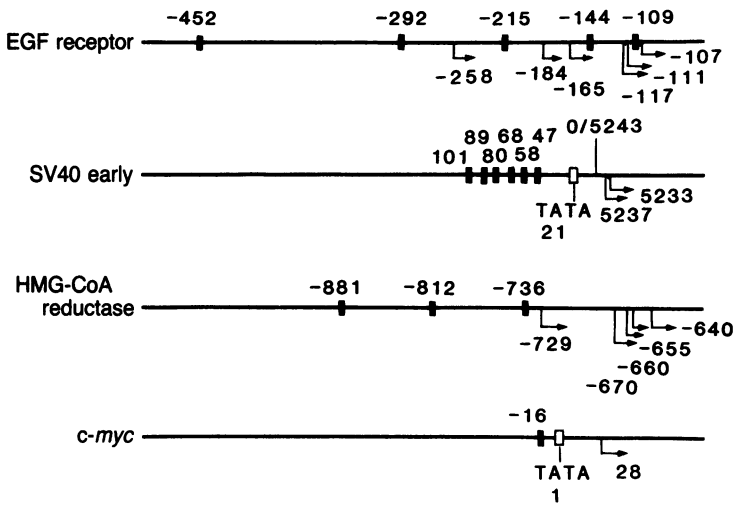

FIG. 6. Diagram of promoter regions containing the CCGCCC sequence. The shaded and unshaded boxes represent CCGCCC sequences and TATA boxes, respectively. The direction and location of the transcriptional initiation sites are indicated by bent arrows. In both the EGF receptor and HMG-CoA reductase (43) genes, A of the initiator methionine is +1. The nucleotide numbers of the SV40 genome are those of the specific SV40 numbering system (44). In the c-*myc* gene (45), the first T of the TATA box is numbered as +1.

specific features of the 5′ flanking region of the EGF receptor gene may prove to be important to our understanding of the regulation of protooncogene expression.

1. Savage, C. R. & Cohen, S. J. (1972) *J. Biol. Chem.* **247**, 7609–7611.
2. Carpenter, G. & Cohen, S. (1979) *Annu. Rev. Biochem.* **48**, 193–216.
3. Ushiro, H. & Cohen, S. (1980) *J. Biol. Chem.* **255**, 8363–8365.
4. Das, M., Miyakawa, T., Fox, C. F., Pruss, R. M., Aharonov, A. & Herschman, H. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2790–2794.
5. Sahyoun, N., Hock, R. A. & Hollenberg, M. D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1675–1679.
6. Beguinot, L., Lyall, R. M., Willingham, M. C. & Pastan, I. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2384–2388.
7. Hunter, T. & Cooper, J. A. (1981) *Cell* **24**, 741–752.
8. Cohen, S., Ushiro, H., Stoscheck, C. & Chinkers, M. (1982) *J. Biol. Chem.* **257**, 1523–1531.
9. Collett, M. S., Purchio, A. F. & Erickson, R. L. (1980) *Nature (London)* **285**, 167–169.
10. Levinson, A. D., Opperman, H., Varmus, H. E. & Bishop, J. M. (1980) *J. Biol. Chem.* **255**, 973–980.
11. Richert, N. D., Blithe, D. L. & Pastan, I. (1982) *J. Biol. Chem.* **257**, 7143–7150.
12. Yamamoto, T., Nishida, T., Miyajima, N., Kawai, S., Ooi, T. & Toyoshima, K. (1983) *Cell* **35**, 71–78.
13. Downward, J., Yarden, Y., Mayes, E., Scarce, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. & Waterfield, M. D. (1984) *Nature (London)* **307**, 521–527.
14. Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D. & Seeburg, P. H. (1984) *Nature (London)* **309**, 418–425.
15. Merlino, G. T., Xu, Y.-h., Ishii, S., Clark, A. J. L., Semba, K., Toyoshima, K., Yamamoto, T. & Pastan, I. (1984) *Science* **224**, 417–419.
16. Xu, Y.-h., Ishii, S., Clark, A. J. L., Sullivan, M., Wilson, R. K., Ma, D. P., Roe, B. A., Merlino, G. T. & Pastan, I. (1984) *Nature (London)* **309**, 806–810.
17. Lin, C. R., Chen, W. S., Kruiger, W., Stolarsky, L. S., Weber, W., Evans, R. M., Verma, I. M., Gill, G. N. & Rosenfeld, M. G. (1984) *Science* **224**, 843–848.
18. Xu, Y.-h., Richert, N., Ito, S., Merlino, G. T. & Pastan, I. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7308–7312.
19. Fabricant, R. N., DeLarco, J. E. & Todaro, G. J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 565–569.
20. Merlino, G. T., Xu, Y.-h., Richert, N., Clark, A. J. L., Ishii, S., Banks-Schlegel, S. & Pastan, I. (1985) *J. Clin. Invest.* **75**, 1077–1079.
21. Libermann, T. A., Nusbaum, H. R., Razon, N., Kris, R., Lax, I., Soreq, H., Whittle, N., Waterfield, M. D., Ullrich, A. & Schlessinger, J. (1985) *Nature (London)* **313**, 144–147.
22. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
23. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
24. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
25. Sanger, R., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
26. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
27. Merlino, G. T., Tyagi, J. S., deCrombrugghe, B. & Pastan, I. (1982) *J. Biol. Chem.* **257**, 7254–7261.
28. Vogeli, G., Ohkubo, H., Sobel, M., Yamada, Y., Pastan, I. & deCrombrugghe, B. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5334–5338.
29. Weaver, R. H. & Weissman, C. (1979) *Nucleic Acids Res.* **7**, 1175–1193.
30. Berk, A. & Sharp, P. (1977) *Cell* **12**, 721–732.
31. Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
32. McKeon, C., Pastan, I. & deCrombrugghe, B. (1984) *Nucleic Acids Res.* **12**, 3491–3502.
33. Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R. & Elgin, S. C. R. (1979) *Cell* **16**, 797–806.
34. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, D. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–695.
35. Weber, W., Gill, G. N. & Spiess, J. (1984) *Science* **224**, 294–297.
36. Merlino, G. T., Ishii, S., Whang-Peng, J., Knutsen, T., Xu, Y.-h., Clark, A. J. L., Stratton, R. H., Wilson, R. K., Ma, D. P., Roe, B. A., Hunts, J. H., Shimizu, N. & Pastan, I. (1985) *Mol. Cell. Biol.*, in press.
37. Banks-Schlegel, S. P. & Harris, C. C. (1983) *Exp. Cell. Res.* **146**, 271–280.
38. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
39. Fromm, M. & Berg, P. (1982) *J. Mol. Appl. Genet.* **1**, 457–481.
40. McKeon, C., Schmidt, A. & deCrombrugghe, B. (1984) *J. Biol. Chem.* **259**, 6636–6640.
41. Elgin, S. C. R. (1981) *Cell* **27**, 413–415.
42. Merlino, G. T., McKeon, C., deCrombrugghe, B. & Pastan, I. (1983) *J. Biol. Chem.* **258**, 10041–10048.
43. Reynolds, G. A., Basu, S. K., Osborne, T. H., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. & Luskey, K. L. (1984) *Cell* **38**, 275–285.
44. Buchman, A., Burnett, L. & Berg, P. (1981) in *DNA Tumor Viruses, Part II Revised*, ed. Tooze, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 799–841.
45. Watt, R., Nishikura, K., Sorrentino, J., ar-Rushdi, A., Croce, C. M. & Rovela, G. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6307–6311.
46. Benoist, C. & Chambon, P. (1981) *Nature (London)* **290**, 304–310.
47. Hansen, U. & Sharp, P. A. (1983) *EMBO J.* **2**, 2293–2303.
48. Wasylyk, B. & Chambon, P. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 921–934.
49. Tjian, R. (1981) *Cell* **26**, 1–2.
50. Gidoni, D., Dynan, W. S. & Tjian, R. (1984) *Nature (London)* **312**, 409–413.
51. Dynan, W. S. & Tjian, R. (1983) *Cell* **32**, 669–680.