Characterization of the Mouse Epidermal Growth Factor Promoter and 5'-Flanking Region

ROLE FOR AN ATYPICAL TATA SEQUENCE*

(Received for publication, July 15, 1996)

Suzanne E. Fenton‡§, Natalie S. Groce‡§, and David C. Lee‡§**

From the ‡Lineberger Comprehensive Cancer Center and §Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina 27599-7295

As a step toward delineating mechanisms that regulate its activity, we have characterized the mouse epidermal growth factor (EGF) promoter. Primer extension and S1 nuclease analyses identified prominent (+1/+2) and minor (+28) transcription start sites, with the dominant +1/+2 site located 33 bases downstream from a TTTAA/A sequence. A restriction fragment that spanned these start sites and contained 390 base pairs of 5'-flanking sequence directed transcription from the +1/+2 site in the presence of HeLa cell nuclear extracts. Additionally, it promoted expression of a coupled luciferase reporter gene in transfected cell lines. The inclusion of additional 5'-flanking sequence either stimulated or inhibited luciferase expression depending on the cell line. Approximately 2 kilobases of EGF 5'-flanking sequence was determined and found to contain several motifs with partial homology to steroid hormone response elements. Despite this fact and evidence that EGF expression might be regulated by androgens in vivo, EGF promoter-luciferase constructs were not steroid-responsive in cells cotransfected with steroid receptor expression vectors. An oligonucleotide containing the aforementioned TTTAA/A sequence specifically bound TATA-binding protein and TFIIA in gel shift assays, and an EGF promoter-luciferase construct in which the core TA dinucleotide was mutated to CG was not active in transfected cells. These data suggest that the TTTAA/A sequence functions as an atypical TATA box.

Epidermal growth factor (EGF)1 was first identified in mouse salivary gland extracts as an activity that induced premature eyelid opening and tooth eruption when injected into newborn pups (1). It was subsequently and independently characterized as urogastrone, a component of human urine that inhibited gastric acid secretion (2). Although its precise physiological roles are still not known, EGF is a potent mitogen for many epithelial and mesenchymal cells, and it regulates cellular migration and differentiation in vitro (3). Its numerous actions are a result of high-affinity binding to the EGF receptor, a type I protein tyrosine kinase (4). EGF is the prototype of the EGF receptor ligand family, which includes the mammalian cellular proteins transforming growth factor-α (5, 6), amphiregulin (7), heparin binding EGF (8), betacellulin (9, 10), and epiregulin (11), as well as several proteins encoded by Shope family viruses (12, 13). Characteristics of this family include a conserved three-loop structure and the proteolytic processing of soluble ligands from the ectodomains of bioactive, integral membrane precursors (reviewed in Ref. 14). The 53-amino acid mature EGF is derived from a precursor protein (prepro-EGF) of approximately 1200 amino acids (15, 16).

The EGF gene is prominently expressed in the granular convoluted tubules of the submaxillary salivary gland (SG), the distal convoluted tubules of the kidney, Brunner’s glands of the duodenum, and alveolar epithelial cells of the mammary gland (17-20). Other sites of expression have been identified, although some remain controversial. Various findings suggest that expression of the EGF gene in SGs may be regulated by androgens. The male mouse SG contains higher levels of EGF mRNA than its female counterpart, and EGF-positive cells contain androgen receptors (17, 21). Moreover, EGF mRNA levels are increased in SGs of female mice given androgen and, conversely, are decreased in SGs of castrated males (22). Other studies suggest that the EGF gene could also be regulated by estrogens (23) and by the lactogenic hormones prolactin and glucocorticoids (24). Finally, deregulation of EGF expression may be a component of neoplastic progression, since EGF mRNA is markedly elevated in some human tumors compared with their normal tissue counterparts (25-30).

The molecular mechanisms that regulate transcription of the EGF gene have not been delineated. As an initial step, we have begun to characterize the EGF promoter. In the present study, we have refined the mapping of transcription start sites, shown the putative promoter to be active in vitro and in vivo, investigated possible regulation by androgens and glucocorticoids, and assessed the role of a TATA-like sequence.

EXPERIMENTAL PROCEDURES

Materials—Luciferin, dexamethasone, and dihydrotestosterone were from Sigma; the 129Sv mouse genomic library and pBluescript SK+ vector were from Stratagene (La Jolla, CA); and radionuclides were from DuPont NEN. The HeLaScribe nuclear extract in vitro transcription system, Altered Sites II in vitro mutagenesis system, pGL2-basic luciferase vector, S1 nuclease, RNasin, and avian myeloblastosis virus reverse transcriptase were from Promega Corp. (Madison, WI). Tissue culture reagents, LipofectAMINE, oligo(dT)-cellulose, Taq polymerase, and DNTPs were from Life Technologies, Inc. Human anti-TFIIId anti-
body was obtained from Santa Cruz Biotechnology (La Jolla, CA). Oligonucleotides were synthesized by the University of North Carolina Nucleic Acids Core Facility.

Recombinant human TATA-binding protein (TBP) and TFIIA were gifts from Robert Roeder (Rockefeller University, New York, NY); the mouse glucocorticoid receptor expression vector was donated by Keith Yamamoto (University of California, San Francisco, CA); and the mouse mammary tumor virus (MMTV)-luciferase construct was obtained from Ron Evans (Salk Institute, San Diego, CA). EGFR probes were gifts from Graeme Bell (University of Chicago, Chicago, IL), and the rat androgen receptor expression vector was provided by Elizabeth Wilson (University of North Carolina at Chapel Hill).

DNA Sequence Analysis—The nucleotide sequence of the EGF HindIII-XhoI fragment was determined from both strands by the University of Wisconsin Automated DNA Sequencing Facility using a dideoxy DNA sequencing (31) was used to resolve ambiguous regions.

Analysis of EGF mRNA 5′-Ends—Cytoplasmic RNA was isolated from cultured cells as described by Gough (32), and poly(A)+ RNA was isolated by oligo(dT)-cellulose chromatography (33). Total RNA was purified from mouse tissues using the guanidinium-cesium chloride method (34). The integrity and concentration of RNAs were verified by gel electrophoresis.

Primer extension and S1 nuclease analyses were performed with 10–20 μg of total RNA as previously reported (35), unless otherwise specified. End-labeled primers were complementary to EGF sequences +131 to +170 (primer 1), +43 to +80 (primer 2), +37 to +76 (primer 3), and +260 to +221 (primer 4) (see Fig. 2A). S1 nuclease probes were generated as previously detailed (35). Primers 1 or 2 were annealed to denatured HX-luc (see Fig. 4A) and extended with the Klenow fragment of E. coli DNA polymerase I. The resulting radiolabeled double stranded products were digested with BamHI (150 bp), and the single stranded probes were isolated from alkaline denaturing gels. Intensities of +1/2 and +28 primer extension and S1 nuclease products were determined by laser scanning densitometry (UltraScan XL; LKB Produkter, Bromma, Sweden).

EGF-Luciferase Constructs—EGF promoter and 5′-flanking restriction fragments possessing a common 3′ end at +314 (XhoI site) were cloned upstream of the luciferase gene. The EGF XX and SX fragments were cloned directly into the corresponding pGL2-basic polylinker sites, whereas the EGF DX fragment was inserted into Smal-XhoI-cleaved pG2L-basic. The EGF HX fragment was excised from XX-XhoI-cleaved Smal-HindIII site in the polylinker and cloned into the HindIII site of the pGL2-basic in the correct orientation (Fig. 4A).

The EGF TTTAAA sequence (−33 to −28 bp) was mutated to a BstBI site (TTCGAAATTAAAAGGAGAG) and the adenosivirus major late promoter TATAAA sequence (−5′GGGTTATAAAAGGAGAG′−3′) were radiolabeled with 32P and separated from free radiolucenode using a 50-column. Purified recombinant TBP (His6-human TBP, 7.1 ng) and TFIIA (8 ng) were added to 20-μl binding reactions containing 20 μm HEPES, pH 7.5, 17 mM KCl, 1 mM dithiothreitol, 0.1 mM EDTA, 4% Ficoll, 0.5 μg poly(dI-dC), 5 μM spermidine, 0.0125% Nonidet P-40, and 75 μM MgSO4. For supershift reactions, the probe, TBP, and TFIIA were first incubated for 15 min on ice, and then 3 μg of human anti-TBP antibody was added for 20 min at room temperature prior to electrophoresis.

RESULTS
Isolation and Characterization of the Mouse EGF Promoter and 5′-Flanking Region—To evaluate the hormonal and tissue-specific regulation of the mouse EGF gene, we isolated genomic sequences containing the EGF promoter and 5′-flanking region from a 129SV mouse liver genomic library. The probe was a 603-bp fragment generated from a mouse kidney genomic DNA via polymerase chain reaction amplification and primers encompassing the putative EGF promoter (36). Eight positive λ-FIXII clones were characterized by restriction enzyme cleavage; the largest (7B) contained approximately 17 kb of sequence 5′ to exon 1 of the EGF gene as well as approximately 4.5 kb of intron 1 sequence (15, 16). The 22-kb NoI fragment from clone 7B was shuttled into pBlueScript SK+ vector, and a partial restriction enzyme map was determined by Southern blotting using the aforementioned 603-bp probe (Fig. 1). Predicted restriction fragments were confirmed by Southern analysis of mouse genomic liver DNA to exclude possible gene rearrangements or splicing events that might have occurred during the cloning process.

Identification of the EGF Transcription Initiation Sites—Using S1 nuclease analysis, a single primer, and mouse SG and kidney RNAs, Pascall and Brown (36) previously identified a single prominent and two minor downstream EGF mRNA 5′-ends. By comparison with a molecular weight ladder, they associated the prominent band with a cytosine residue. Since the mouse EGF promoter contains multiple TA-rich elements that could function as atypical TATA boxes (see Fig. 2A), we wanted to exclude possible transcription initiation at other sites in the flanking regions. Additionally, we wanted to more accurately map transcription initiation sites by comparison with sequence ladders derived from the EGF promoter region itself. Accordingly, we synthesized a set of four oligonucleotide primers, each of which corresponded to sequences downstream of the putative EGF TATAAA and the adenosivirus major late promoter TATAAA sequences. EMSA experiments were performed in duplicate.

Luciferase Expression—Plated cells were resuspended in 1 ml of phosphate-buffered saline (4°C) pelleted at 14,000 rpm, and resuspended in 300 μl of 100 mM K2HPO4, pH 7.8. Cells were then lysed with three successive freeze-thaw cycles, and the luciferase activity of individual cell samples was measured in duplicate. Lysate (100 μl) was placed in the luminometer, and reactions were initiated by automatic injection of 200 μl each of luciferin reagent (200 μM luciferin in 25 mM glycylglycine, pH 7.8) and assay buffer (25 mM glycylglycine, pH 7.8, 15 mM K2HPO4, pH 7.8, 15 mM MgSO4, 4 mM EDTA, 2 mM ATP, and 1 mM dithiothreitol). Luciferase activity was measured for 15 s at ambient temperature imitated by the addition of reagents (AutoLumat LB953 luminometer; Berthold Analytical Instruments, Inc., Nashua, NH). Relative light units were corrected for lysate protein content.

Electrophoretic Mobility Shift Assays—Double stranded DNA probes encoding the EGF promoter TTTAAA sequence (−5′TCGACAGACCTTAAAAGGAGAG−3′) and the adenosivirus major late promoter TATAAA sequence (−5′GGGTTATAAAAGGAGAG−3′) were radiolabeled with 32P and separated from free radiolucenode using a 50-column. Purified recombinant TBP (His6-human TBP, 7.1 ng) and TFIIA (8 ng) were added to 20-μl binding reactions containing 20 μm HEPES, pH 7.5, 17 mM KCl, 1 mM dithiothreitol, 0.1 mM EDTA, 4% Ficoll, 0.5 μg poly(dI-dC), 5 μM spermidine, 0.0125% Nonidet P-40, and 75 μM MgSO4. For supershift reactions, the probe, TBP, and TFIIA were first incubated for 15 min on ice, and then 3 μg of human anti-TBP antibody was added for 20 min at room temperature prior to electrophoresis.

RESULTS

Identification of the EGF Transcription Initiation Sites—Using S1 nuclease analysis, a single primer, and mouse SG and kidney RNAs, Pascall and Brown (36) previously identified a single prominent and two minor downstream EGF mRNA 5′-ends. By comparison with a molecular weight ladder, they associated the prominent band with a cytosine residue. Since the mouse EGF promoter contains multiple TA-rich elements that could function as atypical TATA boxes (see Fig. 2A), we wanted to exclude possible transcription initiation at other sites in the flanking regions. Additionally, we wanted to more accurately map transcription initiation sites by comparison with sequence ladders derived from the EGF promoter region itself. Accordingly, we synthesized a set of four oligonucleotide primers, each of which corresponded to sequences downstream of the putative EGF TATAAA and the adenosivirus major late promoter TATAAA sequences. EMSA experiments were performed in duplicate.

Luciferase Expression—Plated cells were resuspended in 1 ml of phosphate-buffered saline (4°C) pelleted at 14,000 rpm, and resuspended in 300 μl of 100 mM K2HPO4, pH 7.8. Cells were then lysed with three successive freeze-thaw cycles, and the luciferase activity of individual cell samples was measured in duplicate. Lysate (100 μl) was placed in the luminometer, and reactions were initiated by automatic injection of 200 μl each of luciferin reagent (200 μM luciferin in 25 mM glycylglycine, pH 7.8) and assay buffer (25 mM glycylglycine, pH 7.8, 15 mM K2HPO4, pH 7.8, 15 mM MgSO4, 4 mM EDTA, 2 mM ATP, and 1 mM dithiothreitol). Luciferase activity was measured for 15 s at ambient temperature imitated by the addition of reagents (AutoLumat LB953 luminometer; Berthold Analytical Instruments, Inc., Nashua, NH). Relative light units were corrected for lysate protein content.

Electrophoretic Mobility Shift Assays—Double stranded DNA probes encoding the EGF promoter TTTAAA sequence (−5′TCGACAGACCTTAAAAGGAGAG−3′) and the adenosivirus major late promoter TATAAA sequence (−5′GGGTTATAAAAGGAGAG−3′) were radiolabeled with 32P and separated from free radiolucenode using a 50-column. Purified recombinant TBP (His6-human TBP, 7.1 ng) and TFIIA (8 ng) were added to 20-μl binding reactions containing 20 μm HEPES, pH 7.5, 17 mM KCl, 1 mM dithiothreitol, 0.1 mM EDTA, 4% Ficoll, 0.5 μg poly(dI-dC), 5 μM spermidine, 0.0125% Nonidet P-40, and 75 μM MgSO4. For supershift reactions, the probe, TBP, and TFIIA were first incubated for 15 min on ice, and then 3 μg of human anti-TBP antibody was added for 20 min at room temperature prior to electrophoresis.
of three putative TATA box-like sequences as well as the previously reported EGF 5'-end (Fig. 2A). Primers 1–4 were used in primer extension assays, whereas radiolabeled primers 1 and 2 were used to make S1 nuclease probes with uniform 5'-ends produced by BanI cleavage (see Fig. 2A). Since EGF mRNA is expressed at particularly high levels in adult mouse SG and kidney but is present at very low or undetectable levels in brain (17), we used total mRNA from these tissues as templates in the primer extension and S1 nuclease assays.

Fig. 2B shows that when radiolabeled primer 1 was used in primer extension assays with SG and kidney RNAs, a prominent cluster of two or more bands of roughly 170 bases in length was observed. Primer 2 confirmed this result and resolved the cluster to two principle bands; by comparison with an EGF promoter sequencing ladder generated from the same primer, these two bands corresponded to adjacent adenosine residues located immediately 3' to the cytosine previously identified by Pascall and Brown (36). We hereafter refer to the most 5'-adenosine residue as +1. Using RNAs from SG and kidney, primer 1 also detected a less prominent 5'-end corresponding to an adenosine at +28 and occasionally other minor extension products as well. The +28 site likely corresponds to a minor

---

Footnote:

2 S. E. Fenton, unpublished observation.
5'-end previously identified (36) in SG samples, which they associated with an adenosine residue located two bases further upstream. The +28 product could not be confirmed with primer 2, since the latter corresponds to sequences from +43 to +80. However, S1 nuclease assays performed with SG and kidney RNAs and probes generated from primers 1 (Fig. 2B) and 2 (data not shown) yielded prominent products corresponding to both the +1/+2 and +28 sites. In contrast, primer extension reactions with primers 3 and 4 (Fig. 2A) did not yield products with any of the mouse RNAs tested, even when higher concentrations of RNA were used (20 versus 10 μg). The fact that +1/+2 and +28 primer extension and S1 nuclease products were most abundant in SG versus kidney RNA and were not detected with brain samples is consistent with the relative EGF mRNA abundance in these tissues as judged by Northern blot analysis (17). Collectively, these data confirm and refine the previously reported EGF mRNA 5'-ends (36). Specifically, they indicate that transcription of the EGF gene in SG and kidney initiates at two sites, with the +1/+2 site being dominant. Interestingly, densitometric analysis of the autoradiographs shown in Fig. 2B revealed that the +1/+2 site is used 7-fold more frequently than the +28 site in SG, but only 2-fold more frequently in kidney. This suggests that transcription is selectively enhanced via the +1/+2 site in SG.

Transcriptional Activity of the EGF Promoter and 5'-Flanking Region—Functional activity of the putative EGF promoter has not been previously demonstrated. Accordingly, we tested its activity both in vitro and in vivo. To assay its ability to direct transcription in vitro in the presence of crude HeLa cell nuclear extracts (Promega), an XhoI fragment containing 6.7 kb of sequence 5' of the start site was cloned upstream of the firefly luciferase gene (XX-luc), and runoff templates were produced by cleavage at unique sites (BglII or XbaI) within the luciferase gene. The dihydrofolate reductase and CMV promoter constructs were also linearized and used in conjunction with radiolabeled molecular weight markers for size comparisons. As expected, the Scal-linearized dihydrofolate reductase template produced two bands of 780 and 736 bases (37), whereas the linearized CMV template yielded a product of 363 bases (Fig. 3). Transcription of the XbaI-cleaved EGF-luciferase template produced two closely spaced products not observed with the parental luciferase vector. The larger product was a diffuse band(s) of approximately 401–410 bases; the smaller, more distinct product had an estimated length of 385 bases. These sizes are similar to those expected on the basis of nucleotide sequence for runoff products initiated at the +1/+2 (404 bases) and +28 sites (376 bases), respectively (Fig. 3). Consistent with these results, transcription of an EGF-luciferase template that had been cleaved at a BgIII site closer to the EGF promoter produced two similarly spaced bands of appropriately reduced size (data not shown). Additionally, although transcription of alternate EGF-luciferase templates containing 390 (SX-luc) or 2000 (HX-luc) bases of 5'-flanking sequence produced comparable products, a template containing only 30 bases of 5'-flanking sequence (DX-luc) did not yield identifiable transcripts (data not shown). The latter result suggests that sequences upstream of −30 are required for EGF promoter activity in vitro.

To verify correct initiation in vitro, transcripts derived from the XhoI-XhoI EGF-luciferase construct were assayed by primer extension using primer 1. Fig. 3 shows that a dominant extension product corresponding to the +1/+2 site was obtained together with minor, smaller products. Notably, despite the generation of an in vitro transcript the size of which appeared consistent with initiation at +28 (Fig. 2B), no corresponding extension product was observed. However, aberrant minor products were evident at bases +41 and +55.

Transcriptional activity in vivo was established via transient transfection of EGF-luciferase constructs into cultured cell lines. EGF fragments possessing a common 3'-end (XhoI; +314) but containing 6.7 kb (XhoI-XhoI; XX), 2.0 kb (HindIII-XhoI;HX), 0.4 kb (SacI-XhoI; SX), and 0.03 kb (DraI-XhoI; DX) of sequence 5' to the transcriptional start site were cloned upstream of the luciferase reporter gene in pGL2 (Fig. 4A). The final vectors were then transiently transfected into CHO and NRK-52E cells, and the resulting luciferase activity was measured after 48 h. As shown in Fig. 4B, EGF promoter activity was confirmed in both cell types, although the effect on luciferase activity of increasing amounts of EGF 5'-flanking sequence differed considerably between the two cell lines. Thus, in CHO cells, optimal activity was obtained with SX-luc, which yielded a 5-fold increase in luciferase activity relative to the promoterless vector; the inclusion of an additional 5'-flanking sequence decreased luciferase expression from HX- and XX-luc to only 2.2- and 1.3-fold above background, respectively. In contrast, in NRK-52E cells, the larger HX- and XX-luc constructs were most active, yielding 60- and 62-fold increases in luciferase activity over background, respectively. In both CHO and NRK-52E cells, the DX-luc construct, which contained only 30 bases of 5'-flanking sequence, was inactive, consistent with the in vitro transcription results described above. Primer extension assays performed with primer 1 and poly(A)+ RNA from cells transfected with XX-, HX-, and SX-luc yielded extension products that by comparison with an EGF promoter se-
The EGF Promoter and 5′-Flanking Region

The nucleotide sequence of the 2.3-kb HindIII-XhoI fragment, which contains approximately 2 kb of sequence 5′ to the dominant +1/+2 start site, is shown in Fig. 5. The translational start site is at +352 bp (15, 16) and is not shown. The sequence from −897 to +314 bp is identical to that previously reported (36). In addition to a putative atypical TATA box (TTTAAA) at −33 bp (see below), the EGF promoter contains several poly-purine-rich motifs and consensus binding sequences for the transcription factors no. NFκB, GAS, AP-1, AP-2, AP-3, Sp1, p53, and C/EBP (Fig. 5), as defined by the transcription factor data set in the Genetics Computer Group program.

The EGF-Luciferase Constructs Are Not Androgen Responsive—As mentioned above, studies in vitro suggest that the EGF gene may be responsive to androgens. Our sequence revealed that the EGF promoter 5′-flanking sequence from −648 to +314 contains two six-base sequences that are identical to the 3′-portion of the 15 base consensus steroid hormone response element (HRE, GGTACANNNTGTTCT; Ref. 38), and the additional 5′-flanking region from −2048 to −649 includes several other potential half-sites. Furthermore, a 15-base sequence from +226 to +240 is 73% identical to the consensus HRE. To determine whether any of the putative HREs confer direct androgen responsiveness on the promoter, EGF-luciferase constructs were transiently transfected into COS, CHO, and NRK-52E cells either alone or in conjunction with mouse androgen receptor expression vector (provided by Elizabeth Wilson). Following transfection, cells were maintained in serum-free media and after 24 h were exposed to 0.1 nM dihydrotestosterone. An additional 24 h later, cells were harvested, and luciferase expression was measured. For comparison, cells were transfected with an MMTV-luc expression vector (provided by Ron Evans). Fig. 6 shows that activity from MMTV-luc was induced by the combination of androgen receptor expression and dihydrotestosterone treatment in all three cell lines. Relative to expression in nontreated control cells, MMTV-luc activity was increased 4-, 6-, and 9-fold in NRK-52E, CHO, and COS cells, respectively. In contrast, HX-luc activity was decreased in CHO and NRK-52E cells in response to hormone treatment. And although the overall EGF-luc activity was higher in COS cells, it was insignificantly increased in hormone-treated samples (Fig. 6). A similar lack of induction by androgens was observed with the EGF XX- and SX-luc constructs (data not shown). We also tested the EGF promoter for glucocorticoid responsiveness. Whereas MMTV-luc was induced more than 25-fold in the presence of dexamethasone and glucocorticoid receptor (provided by Keith Yamamoto), the activity of the EGF XX-, HX-, and SX-luc constructs was unchanged compared with expression in untreated control cells (data not shown). Interestingly, we note that these experiments revealed the EGF promoter to be dramatically less active in the absence of serum in all three cell lines (compare Figs. 4 and 6).

The Atypical TTTAAA Sequence at −33 bp Is Required for Maximal EGF Promoter Activity—As previously noted, the EGF promoter contains a TTTAAA sequence located from −33
to −27 bp upstream of the +1/+2 start site. The finding that the DX-luc construct, which contains 5′-sequences to −30 bp, had negligible activity in vitro and in vivo is consistent with a possible role for the TTTAAA sequence as an atypical TATA box. To specifically assess the role of the TTTAAA motif, we examined both its ability to bind TBP in vitro and tested its requirement for efficient EGF promoter activity in vivo. To test binding via electrophoretic mobility shift assay, we used a combination of TBP and TFIIA (provided by Robert Roeder), since the binding of TBP to TATA box sequences is facilitated in the presence of TFIIA (39). Fig. 7 shows that a 14-bp probe corresponding to the TATA sequence of the adenovirus major late promoter (AdMLP) displayed the expected mobility shift in the presence of TBP and TFIIA, and that the formation of the product was competitively inhibited in the presence of a 25-fold excess of unlabeled AdMLP double-stranded oligonucleotide. A 20-bp double stranded probe encompassing the EGF promoter TTAAAT sequence displayed a similar mobility shift in the presence of TBP-TFIIA, and this binding was specifically inhibited in the presence of a 25-fold molar excess of unlabeled TTAAAT oligonucleotides. Importantly, the mobility shift was also blocked in the presence of a 25-fold molar excess of the unlabeled AdMLP oligonucleotide, and conversely, the mobility shift of the AdMLP probe was inhibited in the presence of the EGF TTAAAT sequence. In contrast, expression of luciferase from the SX-luc mutant construct was comparable with that of the promoterless vector control in both cell lines. These data confirm that the TTTAAA sequence is required for efficient expression of the EGF promoter in vivo.

**DISCUSSION**

Our results show that transcription of the EGF gene principally initiates at adjacent adenine residues located approximately 30 bp downstream from the TTAAAT sequence, as expected. We compared the activity of a wild-type SX-luc construct with that of a mutant SX-luc in which the TTAAAT sequence had been converted to TTCCGAA by site-directed mutagenesis. Compared with background (pGL2-basic) levels, SX-luc in this experiment yielded 3.7- and 10-fold increases in luciferase expression in CHO and NRK-52E cells, respectively (Fig. 8). In contrast, expression of luciferase from the SX-luc mutant construct was comparable with that of the promoterless vector control in both cell lines. These data confirm that the TTAAAT sequence is required for efficient expression of the EGF promoter in vivo.
inactivity to tissue-specific regulation of the EGF promoter. However, the relationship of the two cell lines to EGF-expressing cells in the distal convoluted tubules in the kidney is uncertain, and the underlying basis of this phenomenon requires further investigation.

Our data indicate that the EGF promoter and 5′-flanking region are not directly responsive to either androgens or glucocorticoids. Work from a number of laboratories suggests that EGF expression can be influenced by these hormones, particularly androgens, in vivo. For example, the SGs of sexually mature male mice were found to contain markedly higher levels of EGF mRNA than those of counterpart females, and the treatment of adult female mice with testosterone resulted in an average 16-fold increase in SG levels of EGF mRNA over a period of several days (22). Similar observations have been made at the protein level. Thus, SGs of male mice contained up to 400 pmol of EGF/mg of protein, whereas corresponding concentrations in female mice were only 5–20 pmol of EGF/mg of protein (40). Moreover, EGF protein levels were increased 4–40-fold in SGs of normal female mice 6 days after administration of testosterone (41), and the corresponding concentrations in androgen-insensitive tfm/y male mice were as low as those of untreated females (42). These various findings have been supported by surgical manipulations; castration at 8 weeks of age resulted in a marked reduction of SG EGF mRNA and protein levels, whereas ovariectomy produced a 100-fold increase in SG EGF mRNA levels (40). Finally, administration of testosterone to hypophysectomized mice induced SG EGF levels nearly 40-fold, with co-administration of testosterone and thyroid hormone producing a synergistic response (43).

In light of the aforementioned observations, we determined the nucleotide sequence of nearly 2500 bp of DNA flanking the transcription start sites to identify potential androgen-responsive elements. The consensus androgen response element GGTCAANNNTGTTCT (38) is similar or identical to the glucocorticoid, progesterone, and mineralocorticoid response elements, and hence the universal term HRE is used. Although the EGF 5′-flanking region does not contain consensus HREs, a 15-bp sequence displaying 73% homology is located approximately 230 bp downstream from the +1/+2 site, and several TGTCTT motifs corresponding to the 3′-6-bp portion of the HRE are present upstream of the start site. Analysis of probasin gene promoter studies indicates that functional androgen-responsive elements can diverge considerably from the 15-bp consensus androgen-responsive element, and that when reiterated, the 3′-TGTCTT sequence can function in the absence of significant homology to the 5′-portion of the HRE (44). Other studies suggest that sequences flanking the putative androgen-responsive element can exert significant influence on hormone responsiveness (45, 46). Hence, it was important to directly test the androgen responsiveness of the EGF promoter. In fact, our data indicate that genomic fragments containing the EGF promoter and up to 7 kb of 5′-flanking sequence are not androgen sensitive. It is still possible that sequences located either far upstream or downstream of the proximal promoter confer androgen responsiveness on the EGF gene. For example, sequences responsible for androgen regulation of the mouse β-glucuronidase gene have been mapped to intron 9 (47). Alternatively, since androgen-induced increases in EGF expression have only been demonstrated in vivo, it is possible that
TTTAAA motif, suggesting that it functions as a cryptic bovine and porcine outer dense fibers (50), prostatic arginine those for the P-450c27/25 (48), herpes simplex virus UL38 (49), upstream of initiation sites in a number of promoters, including

Importantly, our results support a role for the TTTAAA sequence as an atypical TATA box. The TTTAAA sequence is positioned a conventional distance upstream from the transcription start sites. Atypical, but apparently functional, TATA motifs have been described elsewhere in the literature. The TTTAAA sequence is positioned a conventional distance upstream from the TATA box (48). The presence of a functional TATA-like element in the EGF promoter may, at least in part, account for its dramatic expression in mouse SG and kidney. This tissue-specific pattern contrasts with that of the related transforming growth factor α, which is more broadly expressed and at levels significantly lower that those of EGF mRNA in kidney and SG. Interestingly, the transforming growth factor α promoter differs in having a much higher G+C content (>80% versus 45% for the EGF promoter) and multiple binding sites for the transcription factor Sp1 and in not possessing a recognizable TATA-like motif. These are all characteristics of so-called housekeeping gene promoters (reviewed in Ref. 14). A T5C5 sequence in the EGF promoter may, at least in part, account for its regulation of EGF production, as well as the mechanisms by which EGF expression is deregulated in neoplastic progression.

FIG. 8. The TTTAAA sequence is required for EGF promoter activity. EGF promoter-luciferase(SX) vectors containing either the wild-type TTTAAA (~33 to ~28) or the mutant TTCGAA sequences were transfected into CHO and NRK-52E cells as described under “Experimental Procedures.” Luciferase activities obtained with the promoterless vector (pGL2) are shown for comparison. Data are shown as mean ± S.E. (bars); n = 5.

REFERENCES
1. Cohen, S. (1962) J. Biol. Chem. 237, 1555–1562
2. Gregory, H. (1975) Nature 257, 325–327
3. Carpenter, G., and Wahl, M. I. (1990) in Peptide Growth Factors and Their Receptors I (Sporn, M. B., and Roberts, A. B., eds) Springer-Verlag, Berlin
4. Prigent, S. A., and Lemoine, N. R. (1992) Prog. Growth Factor Res. 4, 1–24
5. Todaro, G. J., De Larco, J. E., and Cohen, S. (1976) Nature 260, 23–31
6. Todaro, G. J., and De Larco, J. E. (1978) Cancer Res. 38, 4147–4154
7. Shoyab, M., Pimman, G. D., McDonald, V. L., Bradley, J. G., and Todaro, G. J. (1989) Science 243, 1074–1076
8. Higashiyama, S., Abraham, J. A., Miller, J., Fiddes, J. C., and Klugbrun, M. (1991) Science 251, 936–939
9. Sasada, R., Ono, Y., Taniyama, Y., Shing, Y., Folkman, J., and Igashiri, K. (1993) Biochem. Biophys. Res. Commun. 190, 1173–1179
10. Shing, Y., Christofori, G., Hanahan, D., Ono, Y., Sasada, R., Igashiri, K., and Folkman, J. (1993) Science 259, 1604–1607
11. Toyoda, H., Komurakura, T., Uchida, D., Takayama, Y., Isebe, T., Okuyama, T., and Hanada, K. (1995) J. Biol. Chem. 270, 7495–7500
12. Chang, W., Upton, C., Hue, S.-L., Purchio, A. F., and McFadden, G. (1987) Mol. Cell. Biol. 7, 535–549
13. Chang, W., Macaulay, C., Hue S.-L., Tam, J. P., and McFadden, G. (1990) Virology 179, 926–930
14. Lee, D. C., Fenton, S. E., Berkowitz, E. A., and Hissong, M. A. (1995) Pharmacol. Rev. 47, 51–85
15. Gray, A., Dull, T. J., and Ullrich, A. (1983) Nature 301, 722–725
16. Scott, J., Urdea, M., Quiroga, M., Sanchez-Pescador, R., Fong, N., Selby, M., Rutter, W. J., and Bell, G. I (1983) Science 221, 236–240
17. Rall, L. B., Scott, J., Bell, G. I., Crawford, R. J., Penschow, J. D., Niall, H. D., and Coghlan, J. P. (1985) Differentiation 313, 228–231
18. Saad, E. C., Lakshmnanan, J., Shapiro, L. J., Fischer, D. A., and Barajas, L. (1990) Differentiation 45, 38–43
19. Elder, J. B., Williams, G., Lacey, E., and Gregory, H. (1978) Nature 271, 466–467
20. Sneider, S. M., Brown, C. F., and DeAugustine, R. P. (1991) Proc. Natl. Acad. Sci. U.S.A. 88, 276–280
21. Marrelli, J. J., Gresik, E. W., and Barka, T. (1967) J. Histochem. Cytochem. 15, 1053–1058
22. Guibb, R. M., Shaw, P. A., Gresik, E. W., Onetti-Muda, A., and Barka, T. (1986) Endocrinology 119, 1382–1387
23. DiAugustine, R. P., Petricoin, P., Bell, G. I., Brown, C. F., Korach, K. S., McLaughlan, J. A., and Teng, C. T. (1988) Endocrinology 122, 2355–2363
24. Fenton, S. E., and Shefield, L. G. (1991) Biochem. Biophys. Res. Commun. 181, 1063–1069
25. Yamahas, M., Fujito, T., Ishikawa, T., Shimotosato, T., Yosokazi, H., Yasui, W., and Tahara, E. (1988) Virchows Arch. A Pathol. Anat. Histopathol. 412, 301–306
26. Kobler, M., Jann, I., Wintzer, H.-O., and Baukeht, T. (1989) Anticancer Res. 9, 1537–1548
27. Sugiyama, K., Yonemura, Y., and Miyazaki, I. (1990) Cancer (Phila.) 63, 1557–1561

Acknowldgments—We thank Geri Youngblood, Jan Sumeral, Elizabeth Wilson, and members of the Lee laboratory for constructive advice and Robert Roeder, Ron Evans, Keith Yamamoto, and Elizabeth Wilson for crucial reagents.
28. Barton, C. M., Hall, P. A., Hughes, C. M., Gullick, W. J., and Lemoine, N. R. (1991) *J. Pathol.* **163**, 111–116
29. Isaacs, J. T., Morton, R. A., Martikainen, P., and Isaacs, W. B. (1991) in *Symposium on Growth Factors in Reproduction*. (Schomberg, D. W., ed) Springer-Verlag New York, Inc., New York
30. Fowler, J. E., Jr., Lau, J. L. T., Ghosh, L., Mills, S. E., and Mounzer, A. (1988) *J. Urol.* **139**, 857–861
31. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
32. Gough, N. M. (1988) *Anal. Biochem.* **173**, 93–95
33. Aviv, H., and Leder, P. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 1408–1412
34. Sandgren, E. P., Luetteke, N. C., Palmiter, R. D., Brinster, R. L., and Lee, D. C. (1990) *Cell* **61**, 1121–1135
35. Blasband, A. J., Rogers, K. T., Chen, X., Azizkhan, J. C., and Lee, D. C. (1990) *Mol. Cell. Biol.* **10**, 2111–2121
36. Pascall, J. C., and Brown, K. D. (1988) *J. Mol. Endocrinol.* **1**, 5–11
37. Pierce, A. J., Jambou, R. C., Jensen, D. E., and Azizkhan, J. C. (1992) *Nucleic Acids Res.* **20**, 6583–6587
38. Beato, M. (1989) *Cell* **56**, 335–344
39. Hernandez, N. (1993) *Genes Dev.* **7**, 1291–1308
40. Pascall, J. C., Saunders, J., Blakely, D. M., Laurie, M. S., and Brown, K. D. (1989) *J. Endocrinol.* **121**, 501–506
41. Bullock, L. P., Barthe, P. L., Mowszowicz, I., Orth, D. N., and Bardin, C. W. (1975) *Endocrinology* **97**, 189–195
42. Barthe, P. L., Bullock, L. P., Mowszowicz, I., Bardin, C. W., and Orth, D. N. (1974) *Endocrinology* **95**, 1019–1023
43. Hosoi, K., Maruyama, S., Ueha, T., Sato, S., and Gresik, E. W. (1992) *Endocrinology* **130**, 1044–1055
44. Rennie, P. S., Bruchovsky, N., Leoc, K. J., Sheppard, P. C., McQueen, S. A., Cheng, H., Snoek, R., Hamel, A., Bock, M. E., MacDonald, B. S., Nickel, B. E., Chang, C., Liu, S., Cattini, P. A., and Matusik, E. J. (1993) *Mol. Endocrinol.* **7**, 23–36
45. Roche, P. J., Huare, S. A., and Parker, M. G. (1992) *Mol. Endocrinol.* **6**, 2229–2235
46. Adler, A. J., Scheller, A., and Robins, D. M. (1993) *Mol. Cell. Biol.* **13**, 6326–6335
47. Lund, S. D., Gallagher, P. M., Wang, B., Porter, S. C., and Ganshew, R. E. (1991) *Mol. Cell. Biol.* **11**, 5426–5434
48. Mullick, J., Addya, S., Sucharow, C., and Avadhani, N. G. (1995) *Biochemistry* **34**, 13729–13742
49. Guzowski, J. F., and Wagner, E. K. (1993) *J. Virol.* **67**, 5086–5108
50. Kim, Y., Adham, I. M., Haack, T., Kremling, H., and Engel, W. (1995) *Biol. Chem. Hoppe-Seyler* **376**, 431–435
51. Chadpelaire, P., Gauthier, E., Ho-Kim, M. A., Bissonnette, L., Tremblay, R. R., and Dube, J. Y. (1991) *DNA Cell Biol.* **10**, 49–59
52. Tavianini, M. A., Hayes, T. E., Magazin, M. D., Minth, C. D., and Dixon, J. E. (1984) *J. Biol. Chem.* **259**, 11798–11803
53. Shin, T. H., Paterson, A. J., and Kudlow, J. E. (1995) *Mol. Cell. Biol.* **15**, 4694–4701

The EGF Promoter and 5′-Flanking Region