5736–5741 Nucleic Acids Research, 1993, Vol. 21, No. 24

Chromatin structure of the EGFR gene suggests a role for intron 1 sequences in its regulation in breast cancer cells

Susan A.Chrysogelos
Lombardi Cancer Research Center and the Department of Biochemistry, S-124A, Georgetown University, 3800 Reservoir Road, NW, Washington, DC 20007, USA

Received August 4, 1993; Accepted November 2, 1993

ABSTRACT

The chromatin structure of the epidermal growth factor receptor gene (EGFR) has been analyzed in several human breast cancer cell lines exhibiting a wide range of EGFR expression. Using DNase I, structural differences were identified in the promoter, first exon, and intron 1 of the EGFR gene that correlate with its expression. Specifically, a DNase I hypersensitive site (DH site) around the exon 1/intron 1 boundary occurred preferentially in estrogen receptor positive breast cancer cell lines with low levels of EGFR expression, while a group of DH sites in intron 1 were observed in estrogen receptor negative, high EGFR expressors. Additionally, a region in the promoter was sensitive to DNase I in all breast cancer cells expressing EGFR, but showed differences in both the level of nuclease sensitivity and the extent of the area that was susceptible. Fine mapping by native genomic blotting revealed the presence of multiple protein footprints in both the promoter and first intron of the EGFR gene in MDA-MB-468 cells, a breast cancer cell line that overexpresses the EGFR gene. The appearance of DH sites in intron 1 associated with high levels of EGFR expression suggests that these regions of the gene contain potential enhancer elements, while the absence of a DH site at the exon 1/intron 1 boundary when the gene is up-regulated suggests the action of a repressor that may block transcriptional elongation.

INTRODUCTION

The epidermal growth factor receptor (EGFR) is a transmembrane glycoprotein with tyrosine kinase activity that upon the binding of EGF, or related ligands, transduces a mitogenic signal to the cell (1–3). The extensive homology between EGFR and the avian erythroblastosis *erbB* oncogene strongly suggests that EGFR is the cellular homolog of v-*erbB* (4–7). EGFR is also implicated in cellular transformation since overexpression of the receptor has been shown to result in EGF-dependent transformation of NIH-3T3 cells (8,9), and overexpression of EGFR has been found in a variety of tumors, correlating in several instances with more advanced disease or poor survival (10).

In breast cancer, high levels of EGFR have been found to correlate strongly with poor prognosis, and this overexpression of EGFR is often accompanied by loss of estrogen receptor (ER) (11–18). In general, human tumor cell lines exhibit substantial variation in their level of EGFR (19,20), and the mechanisms responsible for elevated EGFR also differ. Cell lines have been identified with EGFR gene amplification with or without gene rearrangements, as well as overexpression in the absence of gene amplification (20). In human breast cancer cell lines, EGFR expression also varies widely, with gene amplification appearing to be a rare event (21). It has been shown that EGFR messenger RNA levels in breast cancer cell lines correlate with the amount of protein found on the cell surface, and that transcriptional control plays an important role in regulating these differences in expression (21).

The EGFR gene promoter is GC rich, contains no TATA or CAAT box, and has multiple transcription start sites (22). A number of factors have been found to interact with the 5′ flanking region of the EGFR gene including Sp1 (23–26), ETF1, which specifically stimulates *in vitro* transcription of TATA-less promoters (25), TCF, which binds to a region that contains repeats of the sequence 'TCC' (26), and a recently identified repressor, GCF (27). Transient transfection assays in HeLa cells have also identified regions of the promoter and intron 1 that showed enhancer activity, and several binding sites for HeLa cell nuclear factors were found in the intron 1 enhancer (28). However, this sequence did not appear to have enhancer activity in the estrogen-dependent MCF-7 breast cancer cell line.

The chromatin structure of genes has been shown in numerous studies to correlate with their transcriptional activity, and DNase I hypersensitive sites (DH sites) have been found to be associated with important regulatory elements (29). In this study, the pattern of DH sites is analyzed for the EGFR gene in a variety of human breast cancer cell lines (both ER+ and ER−) with a wide range of EGFR expression, and the implications of the results for regions that are functionally involved in the regulation of this gene are discussed.

MATERIALS AND METHODS

Cell lines

The cell lines used in this study were obtained from the following sources: MCF-7, Dr Marvin Rich (Michigan Cancer Foundation, Detroit, MI); MCF-7/ADR, Dr Kenneth Cowan (Clinical Pharmacology Branch, NIH, Bethesda, MD); ZR-75, Dr Marc Lippman (Lombardi Cancer Center, Georgetown University, Washington, DC); T47D, MDA-MB-231, MDA-MB-468, and

MDA-MB-453, American Type Culture Collection (Rockville, MD). All cells were maintained in Richter's modified minimal essential medium (IMEM; Biofluids, Rockville, MD) supplemented with 10% fetal calf serum (Biofluids), and were harvested by trypsinization.

Isolation and treatment of nuclei

Nuclei were isolated by suspension of cells (approximately $1-2\times10^8$ cells per experiment) in 20 ml of reticulocyte standard buffer (RSB; 10mM Tris−HCl, pH 7.4/10mM NaCl/3mM $MgCl_2$) containing 0.5% Nonidet P-40. The nuclei were washed several times in RSB and resuspended at a concentration of approximately 3×10^7 nuclei per ml in RSB. 300μl aliquots of nuclei were digested with DNase I (Sigma) at concentrations ranging from 0.1 to 32 μg/ml for 10 min at 37°C. Controls were an aliquot of nuclei with no DNase I added, kept on ice for 10 min, and an aliquot with no DNase I added, incubated at 37°C for 10 min. Digestion was stopped by the addition of EDTA to 25mM, SDS to 0.5%, and proteinase K to 0.5 mg/ml, followed by incubation at 37°C for 5 hrs.

DNA purification and analysis

DNA was purified by two rounds of extraction with phenol/chloroform, one with chloroform, and ethanol precipitation. The samples were then treated with RNase A, and the extractions and precipitation were repeated. Samples of purified DNA were digested with the appropriate restriction enzyme according to the recommendations of the supplier. For agarose gels, 12 μg of each sample (1μg for MDA-MB-468 samples, which have an amplified EGFR gene) were electrophoresed and alkaline transferred to nylon (Zetabind, AMF/Cuno). Internal marker lanes consisted of HeLa DNA completely digested with the restriction enzyme used for the DNase I treated samples and then partially digested with restriction enzymes for which recognition sites existed in the region of interest in the EGFR gene. Probes were labeled by random oligonucleotide priming, and hybridization was carried out for 20 hrs at 65°C in 5×SSPE (1×SSPE: 180mM NaCl/10 mM$NaPO_4$, pH 8.3, 1mM EDTA)/0.02% Ficoll/0.02% polyvinylpyrrolidone/0.01% BSA/0.2% SDS/100 μg/ml herring testes DNA/10% Dextran Sulfate. After hybridization, filters were washed three times for 30 min with 5mM $NaPO_4$, pH 7.0/1mM EDTA/0.2% SDS at room temperature, once for 30 min with 0.1×SSC (1×SSC: 150mM NaCl/15mM sodium citrate, pH 7.3)/0.1% SDS at 60°C, and once for 30 min with 0.1×SSC/0.1% SDS at 65°C, prior to autoradiography at −70°C with Kodak XAR-5 film and DuPont Lightning Plus intensifying screens.

For the native genomic blots (30,31), 50 μg of HeLa DNA, or 4 to 16 μg of DNase I treated MDA-MB-468 DNA were ethanol precipitated following restriction enzyme digestion, resuspended in TE (10mM Tris−HCl, pH7.5/1mM EDTA), and dried in a speed vac concentrator. Pellets were dissolved in TE and dye mix and electrophoresed through a 4% polyacrylamide gel for 4 hrs at 100V in 0.5×TBE (1×TBE: 100mM Tris, pH 8.3/100mM boric acid/2mM EDTA). The gel was then sealed in a plastic bag and placed in boiling water for 10 min prior to electrotransfer to nylon (GeneScreen, DuPont-NEN) for 90 min at 110V in 0.5×TBE. Following transfer, the DNA samples were UV crosslinked (0.16 kJ/m^2) and probed using a random oligonucleotide primed fragment prepared with four radioactive nucleotides. Hybridization was carried out for 20 hrs at 65°C in 0.5M $NaPO_4$/7% SDS/1mM EDTA and the filters were washed six times for 15 min with 40mM $NaPO_4$/1% SDS/1mM EDTA and exposed to film as described above.

RESULTS

DNase I hypersensitivity of the EGFR gene in two human breast cancer cell lines with different levels of EGFR expression

In order to establish the pattern of DH sites in the EGFR gene in human breast cancer cells, two cell lines were used initially that express very different levels of EGFR. The MCF-7 cell line is ER+ and contains very low amounts of EGFR, while the MCF-7/ADR cell line was derived from MCF-7 by selection for resistance to adriamycin, and has lost ER and increased EGFR expression 100-fold (32). For each of these cell lines, nuclei were isolated and treated with increasing concentrations of DNase I, and the purified DNA was digested with EcoRI and Southern blotted. The 8 kb region that surrounds exon 1 and is defined

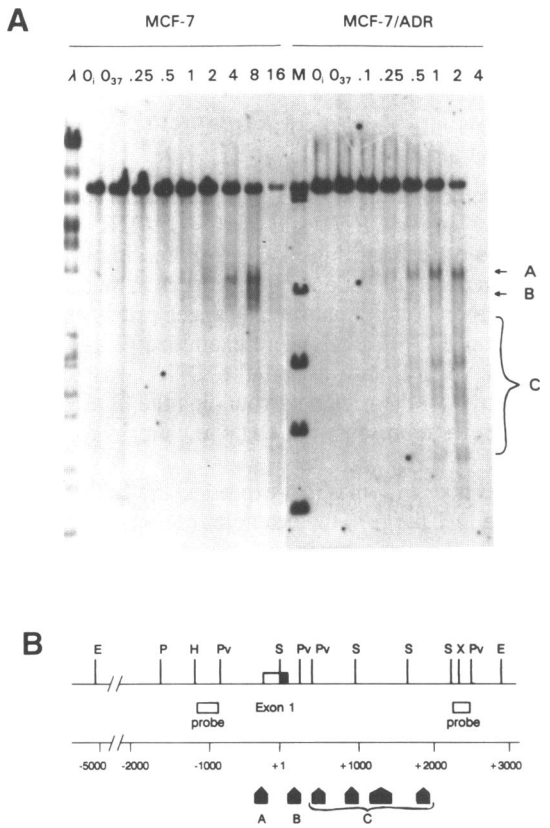

Figure 1. Generation of DH sites in the EGFR gene in breast cancer cells. Panel **A:** Nuclei were isolated from MCF-7 and MCF-7/ADR cells, and treated with the indicated concentration of DNase I. The purified DNA was digested with EcoRI, Southern blotted, and probed with the intron 1 fragment indicated in panel B. The DNase I-generated sub-bands A, B and C correspond to the DH sites shown in panel B. Lanes from left to right are: λ = λHindIII + λEcoRI/HindIII markers; M = HeLa DNA partially digested with SstI (marker lane); 0_i = control nuclei, no DNase I added, kept on ice; 0_{37} = control nuclei, no DNase I added, 37°C, 10 min.; 0.1−16 = concentration of DNase I (μg/ml) used to treat nuclei, 37°C, 10 min. Panel **B:** Map of the 8 kb EcoRI fragment surrounding exon 1 of the EGFR gene. Position +1 is the translational start site, and the probes used in panel A and Figure 2 are indicated by the open boxes. The position of the DH sites indicated in panel A are designated by the solid arrows marked A, B and C. Restriction sites: E, EcoRI; P, PstI; H, HindIII; Pv, PvuII; S, SstI; X, XbaI.

by EcoRI sites (Fig. 1B) was then probed by indirect end-labeling using a small fragment from intron 1. A lane of HeLa DNA partially digested with SstI provided the locations of the SstI sites within the EcoRI fragment as internal markers.

Several differences in the DNase I sensitivity of this region of the EGFR gene were apparent in the two cell lines (Fig. 1A). First, it was observed that the overall level of DNase I sensitivity for the EGFR gene was approximately four to eight fold higher in the MCF-7/ADR cells than in the MCF-7 cells (compare MCF-7 lanes 8 and 16 with MCF-7/ADR lanes 2 and 4), consistent with the increased level of EGFR expression in the MCF-7/ADR cells. Second, distinct patterns of DH sites were found for the two cell lines. Most notably, a group of sites (C) was observed in the MCF-7/ADR cells that was located in the EGFR first intron between nucleotides +500 and +2000, and was not found in the MCF-7 cells. Additionally, a DH site (B) was seen in the MCF-7 cells that was absent in the MCF-7/ADR cells, and this site was mapped to the exon 1/intron 1 boundary. A DH site in the promoter region of the EGFR gene just upstream from the major start site of transcription (site A) was present in both cell lines. Though the EcoRI digestion used for this experiment permitted the analysis of 5 kb of the EGFR promoter, no other DH sites were observed 5′ to the EGFR gene.

Consensus DH site patterns for the EGFR gene in human breast cancer cell lines

In order to determine if the DH site patterns observed for the EGFR gene in MCF-7 and MCF-7/ADR cells were representative of low and high EGFR expressing breast cancer cell lines in general, or reflected differences in cell line specific regulation, several additional breast cancer cell lines exhibiting a wide range of EGFR levels were examined (Table 1). Nuclei were treated with DNase I and the DNA was analyzed as described above for the region surrounding the EGFR first exon. Figure 2A shows representative DH site patterns for the EGFR gene in each of these cell lines.

All breast cancer cell lines that expressed the EGFR gene were found to have DH site A in the promoter region, though the level of sensitivity in this region appeared to vary among the cell lines. The three ER+ cell lines that have low to moderate levels of EGFR all showed DH site B, which could be seen in this experiment to consist of two bands, one of which mapped to the exon 1/intron 1 boundary and the other 150 bp into the first intron. The three ER− cell lines that overexpress EGFR all contained the group of DH sites (C) in intron 1. Additionally, DH site B appeared to be lacking or greatly reduced in intensity in these cell lines. Conversely, the ER+ cell lines lacked the intron 1 DH sites with the possible exception of the T47D cells, which contain higher levels of EGFR than the MCF-7 and ZR-75 cells, and show a faint suggestion of DH sites in the first intron. The faint band seen 5′ to DH site A in the T47D and MDA-MB-231 cells was occasionally observed in other experiments and in different cell lines. However it appeared independent of

Table 1. EGFR and ER levels in breast cancer cell lines

Cell line	EGFR[a]	ER[a]
MCF-7	+/−	+
ZR-75	+/−	+
T47D	+	+
MDA-MB-231	++	−
MCF-7/ADR	++[b]	−[b]
MDA-MB-468	+++[c]	−[c]
MDA-MB-453	−[d]	−[d]

[a] Data from ref. 21, except where noted. EGFR: −, no expression; +/−, low level expression (<10,000 sites per cell); +, moderate expression (10,000−70,000 sites per cell); ++, overexpression (>70,000 sites per cell); +++, gene amplification and overexpression. ER: +, positive; −, negative.
[b] Data from ref. 32.
[c] Data from ref. 33.
[d] Data from ref. 34.

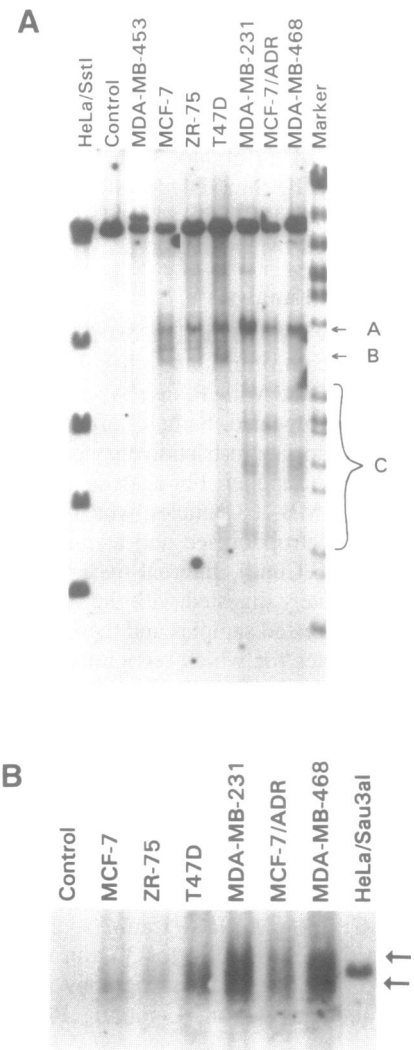

Figure 2. Comparison of DH sites in ER+ and ER− breast cancer cell lines. Nuclei were isolated from the indicated cell lines and treated with DNase I. Panel **A**: The purified DNA was digested with EcoRI, Southern blotted, and probed with the intron 1 fragment indicated in Figure 1B. The DNase I-generated sub-bands A, B and C correspond to the DH sites shown in Figure 1B. Lanes from left to right are: HeLa DNA partially digested with SstI (marker lane); control nuclei from MCF-7/ADR cells, untreated with DNase I; MDA-MB-453 cells; MCF-7 cells; ZR-75 cells; T47D cells; MDA-MB-231 cells; MCF-7/ADR cells; MDA-MB-468 cells; and λ HindIII + λ EcoRI/HindIII markers. Panel **B:** Resolution of DH site A. The purified DNA was digested with HindIII and XbaI, Southern blotted, and probed with the promoter fragment indicated in Figure 1B. The arrows indicate the 5′ (lower) and 3′ (upper) halves of DH site A. Lanes from left to right are: control nuclei from MCF-7/ADR cells, untreated with DNase I; MCF-7 cells; ZR-75 cells; T47D cells; MDA-MB-231 cells; MCF-7/ADR cells; MDA-MB-468 cells; and HeLa DNA partially digested with Sau3AI (marker lane).

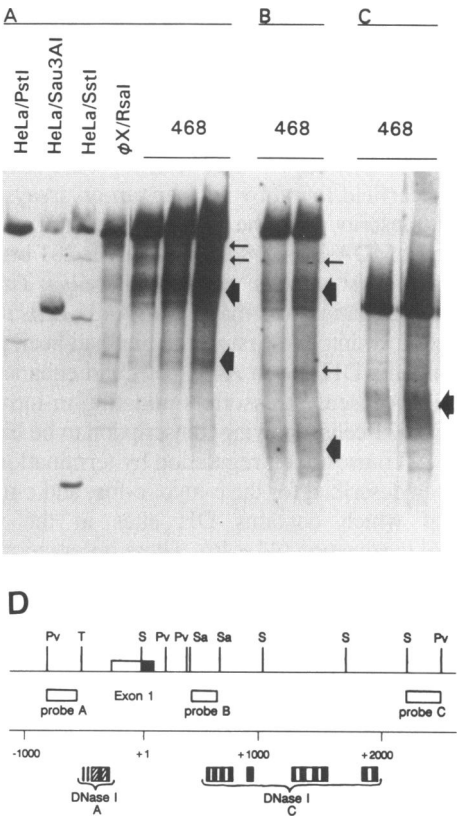

Figure 3. Fine mapping of DNase I hypersensitive and protected sites in the EGFR gene by native genomic blotting. Nuclei were isolated from MDA-MB-468 cells and treated with DNase I. The purified DNA was digested with PvuII, electrophoresed through a native 4% polyacrylamide gel, electrotransferred to a nylon membrane, and probed with the fragments from the EGFR gene indicated in panel D. Markers are digests of φX DNA and partial restriction digests of HeLa DNA. Panel **A:** Lanes from left to right are: HeLa/PstI digest; HeLa/Sau3AI digest; HeLa/SstI digest; φX/RsaI with 50μg carrier control genomic DNA; 4, 8, and 16μg DNA from DNase I treated MDA-MB-468 nuclei. The probe is a 3′ intron 1 fragment, and the arrows correspond to DH site C from 5′ to 3′ (top to bottom). Panel **B:** MDA-MB-468 DNase I treated samples reprobed with a 5′ intron 1 fragment. The arrows correspond to DH site C from 3′ to 5′ (top to bottom). Panel **C:** MDA-MB-468 DNase I treated samples reprobed with a promoter fragment. The arrow indicates DH site A. Panel **D:** Map of the region surrounding exon 1 of the EGFR gene. Position +1 is the translational start site, and the probes used in panels A, B, and C are indicated by open boxes. The positions of the DH sites and footprints seen in panels A, B, and C are designated by the series of closed and open boxes (respectively) at the bottom of the map. These sites are labeled DNase I A and DNase I C, which correspond to the DH sites A and C shown in Figures 1 and 2. The hatched region in the DNase I A box corresponds to the broad nuclease sensitive region in the promoter described in the text. Restriction sites: Pv, PvuII; T, TaqI; S, SstI; Sa, Sau3AI.

DNase I, was very resistant to digestion, and varied with wash conditions, suggesting that it represents nonspecific cross-hybridization rather than an additional DH site.

The MDA-MB-453 cell line does not express EGFR (34), and as seen in Figure 2A, no specific DH sites were found for the EGFR gene in these cells. Interestingly, two different size alleles were noted for EGFR in this cell line and they showed different sensitivities to DNase I. The normal 8 kb EcoRI fragment detected with this probe was very resistant to DNase I digestion, showing the same kinetics of digestion as bulk DNA. The abnormal allele resulted in a slightly larger EcoRI fragment that displayed DNase I sensitivity similar to the low expressing MCF-7 cells. In general, the cell lines followed the pattern set by the MCF-7 and MCF-7/ADR cells, with the high EGFR expressors showing a greater degree of sensitivity to DNase I than the low expressors (data not shown).

Because of the variability in DH site A from cell line to cell line, an attempt was made to resolve this region at higher resolution. Purified DNA from DNase I treated nuclei was digested with HindIII and XbaI and probed with a fragment from the EGFR promoter (see Figure 1B). Figure 2B shows DH site A for the three ER+ and three ER− breast cancer cell lines at this higher resolution. The three high EGFR expressing, ER− cell lines appear to have a doublet for DH site A, while the ER+, low expressors have only the more 5′ half of the site. The position of a Sau3AI site from the internal marker lane places the 5′ portion of DH site A over the more upstream TCC repeats in the EGFR promoter, while the 3′ half of the site covers the area that includes the more downstream TCC repeats.

Fine mapping of protein−DNA binding sites in a breast cancer cell line that overexpresses EGFR

To identify specific binding sites for regulatory factors for the EGFR gene, it is necessary to increase resolution sufficiently to detect protein footprints within the DH sites. This fine mapping was performed using the native genomic blotting technique, which permits the visualization of both DH sites (as dark bands) and protein footprints (as light, protected bands) in one gel (30,31). This method is based on the *in vivo* DNase I footprinting/genomic sequencing procedures, however the DNA samples are electrophoresed through a small, non-denaturing polyacrylamide gel, and the DNA is then denatured by boiling the gel prior to electrotransfer to a nylon membrane.

The MDA-MB-468 breast cancer cell line was used for these experiments since it contains an EGFR gene that is amplified approximately 16 fold (33), thus improving detection considerably. Purified DNA from DNase I treated nuclei was digested with PvuII, electrophoresed through a native 4% polyacrylamide gel, electrotransferred to a nylon membrane, and probed with the fragments from the EGFR gene indicated in Figure 3D. Size markers were provided by an end-labeled RsaI digest of φX DNA, and by partial restriction digests of HeLa DNA, which serve as internal markers.

The probe used in Figure 3A is a 250 bp SstI/PvuII fragment from the 3′ end of intron 1. From this experiment it is evident that the broad, fuzzy sub-bands seen in the traditional Southern blot in Figure 2A can now be resolved to show the presence of footprints within them. Since the best resolution with native genomic blotting is obtained in the size range of 200 to 1000 bp, and the intron 1 PvuII fragment is 2200 bp, a 310 bp Sau3AI fragment from the 5′ end of intron 1 was used as an additional probe as shown in Figure 3B. Because this probe is from the opposite end of the PvuII restriction fragment, the order of the sub-bands is reversed from that in 3A, and the structure of the more 5′ portion of intron 1 can be better resolved. Finally, a 210 bp PvuII/TaqI fragment from the promoter was used to probe a 1000 bp PvuII fragment that contains exon 1 and a portion of the promoter (Fig. 3C).

To determine that the footprints identified by native genomic blotting were not due to sequence specific DNase I digestion, purified genomic DNA was cut with PvuII, partially digested with DNase I, and analyzed by native genomic blotting using the EGFR probes described above. No specific pattern of DNase

I digestion was observed, confirming that the protected bands are actually protein footprints (data not shown).

The native genomic blotting results for the MDA-MB-468 cell line are summarized in Figure 3D, which indicates the position of DH sites in the EGFR gene, and locations where protected sites were found within them. The spacing of the DNase I sensitive and protected sites in intron 1 clearly indicates that nuclease sensitivity cannot be due simply to nucleosomal phasing, and that specific regions of the intron (particularly from approximately +600 to +750 and from +1300 to +1550) are involved in the binding of multiple factors. In the promoter, a broad DNase I sensitive region was seen to extend to the promoter/exon 1 junction and contains hypersensitive sites within it in the TCC repeat region. Additionally, more 5′, minor DH sites were also detected.

DISCUSSION

The results of this study indicate that there are differences in the chromatin structure of the EGFR promoter, first exon, and intron 1 that correlate with the expression of EGFR in breast cancer cell lines. The pattern of DNase I sensitivity in seven cell lines (three ER+, low EGFR expressors; three ER−, EGFR overexpressors; and one line with no EGFR expression) was analyzed and the following consensus DH sites were observed. Breast cancer cell lines that overexpressed EGFR contained a group of DH sites in intron 1, while the low EGFR expressing cell lines had a DH site at the exon 1/intron 1 boundary. All the cell lines expressing EGFR were found to have a DH site in the promoter just upstream from the major start site of transcription, however this site showed cell line variability in the both its size and level of nuclease sensitivity. No DH sites were found when the EGFR gene was not expressed. Additionally, the degree of DNase I sensitivity of the EGFR gene correlated in general with the level of EGFR expression for the different cell lines.

The appearance of DH sites in intron 1 associated with overexpression of EGFR indicates that these regions of the gene are probably involved in the binding of positive regulatory factors, and fine mapping by native genomic blotting confirmed the presence of multiple protein footprints within these DH sites in a breast cancer cell line overexpressing EGFR. There is a growing body of evidence documenting the importance of intronic elements, particularly in the first intron of genes, in transcriptional regulation. Several genes, including human PDGF-B, collagen alpha (I) and IFN (gamma), mouse HPRT and TIMP, and Drosophila beta 3-tubulin, contain enhancer elements in their first introns (35−40). Some of these elements were found to have nonspecific, general enhancer activity (37), while others showed tissue specificity (36, 40) or functioned only in collaboration with the homologous promoter (35). The PDGF-B gene is of particular interest since it was found to contain a series of cell type-specific DH sites in its first intron that identified sequence elements conferring cell type-specific transcriptional regulation on its own promoter (35). Experiments are currently in progress to define the regulatory functions of the sequence elements that comprise the EGFR intron 1 DH sites.

The preferential occurrence of a DH site at the exon 1/intron 1 boundary in cell lines with low levels of EGFR suggests the action of a repressor that may block transcriptional elongation. Alternatively, since the cell lines that show the presence of this DH site are ER+ and were maintained in the presence of phenol red (a component of which can produce estrogenic effects [41]) and fetal calf serum (which contains estrogenic compounds [42]), it is possible that this site is due to the binding of factors in response to estrogen. Evidence for a block to transcriptional elongation in the EGFR gene has been reported recently (43). Haley and Waterfield found by nuclear run-on assays that there was elevated transcription of the EGFR first exon in A431 vulval carcinoma cells, MDA-MB-468 and MDA-MB-231 breast cancer cells, and HN-5 squamous carcinoma cells. The site of transcription termination was mapped in A431 cells to a region approximately 2 kb into the first intron that had been previously shown to contain a DH site in A431 cells and enhancer activity in HeLa cells. Antisense transcripts initiating in intron 1 were also found in A431 cells, showing transcription to be bidirectional across exon 1. Transcription regulation by termination in intron 1 has also been described for the c-myc, c-fos, and c-myb genes, the latter of which contains DH sites in the region of transcriptional termination (44−46). These observations raise the possibility that the different levels of EGFR found in ER+ and ER− breast cancer cells may, in part, be controlled by a block in transcriptional elongation that varies in its degree.

The DH site in the EGFR promoter is located in an area that contains four repeats, each 10 to 15 base pairs in length, of the sequence 'TCC'. The presence of the 5′ half of this DH site in all breast cancer cell lines examined that express EGFR implicates these sequences in the regulation of a basal level of EGFR expression, while the 3′ half of this site appears to be specific for cells overexpressing EGFR. The 5′ half of this region has also been shown to bind a nuclear factor present in KB epidermoid carcinoma cells, and to be sensitive to the single-strand specific nuclease, S1 (47−48). This S1 sensitivity is typical of homopurine/homopyrimidine stretches, and there has been much speculation as to the structure of these regions. One theory is that these sequences respond to the surrounding chromatin structure by 'sliding' of the two DNA strands relative to each other to produce single-stranded loops on both the coding and non-coding stands, while another proposes a folding back of the DNA to yield a triple helix structure and a single-stranded region on only one strand (47,49). Different factors present in cell lines with low and high levels of EGFR may be involved in determining the three dimensional structure of this region of DNA, which may in turn control interactions between the promoter and other more distant regulatory elements.

In a study characterizing the EGFR promoter in A431 vulval carcinoma cells (a cell line with an amplified, overexpressed EGFR gene), Ishii *et al.* (22) found two DH sites for the EGFR gene. One site mapped to the TCC repeat region of the promoter, while the other was localized to a position approximately 2 kb into the first intron. The finding of a DH site in the same region of the promoter in both breast cancer cells and vulval carcinoma cells again implicates these sequence elements in basal, nonspecific regulation of the EGFR gene. However, the different DH site patterns seen in intron 1 in these two cell types suggests that this region may contain elements responsible for tissue-specific enhanced expression of EGFR. In support of this hypothesis, the same DH site pattern that was found in breast cancer cells overexpressing EGFR was also observed in 184A1N4 cells (unpublished results), a derivative of the normal human mammary epithelial line 184 that was immortalized by benzo-a-pyrene, is non-tumorigenic, and overexpresses EGFR (50, 51).

Maekawa *et al.* (28) also found enhancer activity in HeLa cells for the intron 1 region that contained a DH site in A431 cells, and the region essential for enhancer activity was shown to have 10 binding sites for HeLa nuclear factors. Of the intron 1 DH sites observed in breast cancer cells, only the most 3′ may possibly involve a portion of the enhancer region identified in HeLa cells, and this site does not appear to overlap with the essential portion for enhancer activity in HeLa cells, again implying that the intron 1 elements are involved in tissue-specific regulation of EGFR. The absence of the intron 1 DH sites in MCF-7 breast cancer cells also agrees with the lack of activity for the HeLa enhancer element that was observed by Maekawa *et al.* in this cell line.

The complex patterns of DH sites observed in the EGFR gene in breast cancer cell lines expressing different levels of EGFR suggest that regulation of this gene may involve a number of interactions between elements within intron 1, or between elements in the promoter and intron 1. The likelihood that factors which interact with intron 1 elements may be tissue-specific has significant implications since the acquisition of EGFR over-expression appears to be an important step in the progression of breast cancer to estrogen independence.

ACKNOWLEDGEMENTS

I thank Drs Glenn Merlino and Shunsuke Ishii for providing the EGFR clones used to generate probes, Dr Marc Lippman for critical reading of the manuscript, Drs Michael Johnson and Dorraya El-Ashry for many helpful discussions, and Roberto Cuca for assistance with tissue culture. This work was supported in part by a grant from the National Cancer Institute (CA55677), and the Nina Hyde Center for Breast Cancer Research.

REFERENCES

1. Cohen,S., Ushiro,H., Stoscheck,C. and Chinkers,M. (1982) *J. Biol. Chem.* 257, 1523–1531.
2. Ushiro,H. and Cohen,S. (1980) *J. Biol. Chem.* 255, 8363–8365.
3. Hunter,T. and Cooper,J.A. (1981) *Cell* 24, 741–752.
4. Downward,J., Yarden,Y., Mayes,E., Scarce,G., Totty,N., Stockwell,P., Ullrich,A., Schlessinger,J. and Waterfield,M.D. (1984) *Nature* 307, 521–2527.
5. Xu,Y.-H., Ishii,S., Clark,A.J.L., Sullivan,M., Wilson,R.K., Ma,D.P., Row,B.A., Merlino,G.T. and Pastan,I. (1984) *Nature* 309, 806–810.
6. 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. and Seeburg,P.H. (1984) *Nature* 309, 418–425.
7. Lin,C.R., Chen,W.S., Verma,I.M., Gill,G.N. and Rosenfeld,M.G. (1984) *Science* 224, 843–848.
8. Velu,T.J., Beguinot,L., Vass,W.C., Willingham,M.C., Merlino,G.T., Pastan,I. and Lowry,D.R. (1987) *Science* 238, 1408–1410.
9. DiFiore,P.P., Pierce,J.H., Fleming,T.P., Hazon,R., Ullrich,A., King,C.R., Schlessinger,J. and Aaronson,S.A. (1987) *Cell* 51, 1063–1070.
10. Gullick,W.J. (1991) *British Med. Bulletin* 47, 87–98.
11. Fitzpatrick,S.L., Brightwell,J., Wittliff,J., Barrows,G.H. and Schultz,G.S. (1984) *Cancer Res.* 44, 3448–3453.
12. Sainsbury,J.R.C., Farndon,J.R., Sherbert,G.V. and Harris,A.L. (1985) *Lancet* i, 364–366.
13. Nicholson,S., Sainsbury,J.R.C., Halcrow,P., Chambers,P., Farndon,J.R. and Harris,A.L. (1989) *Lancet* i, 182–184.
14. Nicholson,S., Halcrow,P., Sainsbury,J.R.C., Angus,B., Chambers,P., Farndon,J.R. and Harris,A.L. (1988) *Br. J. Cancer* 58, 810–814.
15. Nicholson,S., Richard,J., Sainsbury,C., Halcrow,P., Kelly,P., Angus,B., Wright,C., Henry,J., Farndon,J.R. and Harris,A.L. (1991) *Br. J. Cancer* 63, 146–150.
16. Bolufer,P., Miralles,F., Rodriguez,A., Vasquez,C., Lluch,A., Garcia-Conde,J. and Olmos,T. (1990) *Eur. J. Cancer* 26, 283–290.
17. Lewis,S., Locker,A., Todd,J.H., Bell,J.A., Nicholson,R., Elston,C.W., Blamey,R.W. and Ellis,O. (1990) *J. Clin. Pathol.* 43, 385–389.
18. Toi,M., Osaki,A., Yamada,H. and Toge,T. (1991) *Eur. J. Cancer* 27, 977–980.
19. Xu,Y.H., Richert,N., Ito,S., Merlino,G.T. and Pastan,I. (1984) *Proc. Natl. Acad. Sci., USA* 81, 7308–7312.
20. King,C.R., Kraus,M.H., Williams,L.T., Merlino,G.T., Pastan,I.H. and Aaronson,S.A. (1985) *Nucl. Acids Res.* 13, 8477–8486.
21. Davidson,N.E., Gelmann,E.P., Lippman,M.E. and Dickson,R.B. (1987) *Mol. Endocrinol.* 1, 216–223.
22. Ishii,S., Xu,Y.-H., Stratton,R.H., Roe,B.A., Merlino,G.T. and Pastan,I. (1985) *Proc. Natl. Acad. Sci., USA* 82, 4920–4924.
23. Johnson,A.C., Ishii,S., Jinno,Y., Pastan,I. and Merlino,G.T. (1988) *J. Biol. Chem.* 263, 5693–5699.
24. Kageyama,R., Merlino,G.T. and Pastan,I. (1988) *J.Biol. Chem.* 263, 6329–6336.
25. Kageyama,R., Merlino,G.T. and Pastan,I. (1988) *Proc. Natl. Acad. Sci., USA* 85, 5016–5020.
26. Merlino,G.T., Johnson,A.C., Kageyama,R. and Pastan,I. (1989) In Lippman,M.E. and Dickson,R.B. (eds.), Growth Regulation of Cancer II: UCLA Symposia on Molecular and Cellular Biology. Alan R. Liss, Inc., New York, Vol. 115.
27. Kageyama,R. and Pastan,I. (1989) *Cell* 59, 815–825.
28. Maekawa,T., Imamoto,F., Merlino,G.T., Pastan,I. and Ishii,S. (1989) *J. Biol. Chem.* 264, 5488–5494.
29. Gross,D.S. and Garrard,W.T. (1988) *Ann. Rev. Biochem.* 57, 159–197.
30. Pauli,U., Chrysogelos,S., Stein,J. and Stein,G. (1988) *Proc. Natl. Acad. Sci., USA* 85, 16–20.
31. Pauli,U., Chrysogelos,S., Stein,J. and Stein,G. (1988) *BioTechniques* 6, 142–146.
32. Vickers,P.J., Dickson,R.B., Shoemaker,R. and Cowan,K.H. (1988) *Mol. Endocrinol.* 2, 886–892.
33. Filmus,J., Pollak,M.N., Cailleau,R. and Buick,R.N. (1985) *Biochem. Biophys. Res. Commun.* 128, 898–905.
34. Kraus,M.H., Popescu,N.C., Amsbaugh,S.C. and King,C.R. (1984) *EMBO J.* 6, 605–610.
35. Franklin,G.C., Donovan,M., Adam,G.I.R., Holmgren,L., Pfeifer-Ohlsson,S. and Ohlsson,R. (1991) *EMBO J.* 10, 1365–1373.
36. Bornstein,P., McKay,J., Liska,D.J., Apone,S. and Devarayalu,S. (1988) *Mol.Cell.Biol.* 8, 4851–4857.
37. Sica,A., Tan,T.H., Rice,N., Kretzschmar,M., Ghosh,P. and Young,H.A. (1992) *Proc. Natl. Acad. Sci., USA* 89, 1740–1744.
38. Magin,T.M., McEwan,C., Milne,M., Pow,A.M., Selfridge,J. and Melton,D.W. (1992) *Gene* 122, 289–296.
39. Coulombe,B., Ponton,A., Daigneault,L., Williams,B.R. and Skup,D. (1988) *Mol. Cell. Biol.* 8, 3227–3234.
40. Gasch,A., Hinz,U. and Renkawitz-Pohl,R. (1989) *Proc. Natl. Acad. Sci., USA* 86, 3215–3218.
41. Berthois,Y., Katzenellenbogen,J.A. and Katzenellenbogen,B.S. (1986) *Proc. Natl. Acad. Sci, USA* 83, 2496–2500.
42. Page,M.J., Field,J.K., Everett,N.P. and Green,C.D. (1983) *Cancer Research* 43, 1244–1249.
43. Haley,J.D. and Waterfield,M.D. (1991) *J. Biol. Chem.* 266, 1746–1753.
44. Bentley,D.L. and Groudine,M. (1986) *Nature* 321, 702–706.
45. Mechti,N., Piechaczyk,M., Blanchard,J.M., Jeanteur,P. and Lebleu,B. (1991) *Mol. Cell. Biol.* 11, 2832–2841.
46. Bender,T.P., Thompson,C.B. and Kuehl,W.M. (1987) *Science* 237, 1473–1476.
47. Kato,M., Kudoh,J. and Shimizu,N. (1990) *Biochem. J.* 268, 175–180.
48. Johnson,A.C., Jinno,Y. and Merlino,G.T. (1988) *Mol. Cell. Biol.* 8, 4174–4184.
49. Johnston,B.H. (1988) *Science* 241, 1800–1804.
50. Clark,R., Stampfer,M.R., Milley,R., O'Rourke,E., Walen,K.H., Kriegler,M., Kopplin,J. and McCormick,F. (1988) *Cancer Res.* 48, 4689–4694.
51. Valverius,E.M., Bates,S.E., Stampfer,M.R., Clark,R., McCormick,F., Salomom,D.S., Lippman,M.E. and Dickson,R.B. (1989) *Mol.Endocrinol.* 3, 203–214.