## Structural alterations of the epidermal growth factor receptor gene in human gliomas

(oncogene/amplification/rearrangement/erbB/tumor-specific surface molecule)

Albert J. Wong\*<sup>†</sup>, John M. Ruppert\*, Sandra H. Bigner<sup>‡</sup>, Carl H. Grzeschik<sup>§</sup>, Peter A. Humphrey<sup>‡</sup>, Darrell S. Bigner<sup>‡</sup>, and Bert Vogelstein\*<sup>¶</sup>

\*The Oncology Center, The Johns Hopkins University School of Medicine, 424 North Bond Street, Baltimore, MD 21231; <sup>‡</sup>Department of Pathology, Duke University Medical Center, Durham, NC 27710; and <sup>§</sup>Institut für Humangenetik, Universität Münster, Münster, Federal Republic of Germany

Communicated by Alfred G. Knudson, December 30, 1991 (received for review December 21, 1990)

ABSTRACT The epidermal growth factor receptor (EGFR) gene is amplified in 40% of malignant gliomas, and the amplified genes are frequently rearranged. We have characterized the genetic alterations associated with these rearrangements in five malignant gliomas. In one tumor the rearrangement resulted in the deletion of most of the extracytoplasmic domain of the receptor, resulting in a hybrid mRNA between new sequences and the truncated EGFR sequence. The predicted amino acid sequence of the protein from this tumor was remarkably similar to that described for several viral erbB oncogenes. Four other tumors were noted to have internal deletions of the EGFR gene. These rearrangements brought about in-frame deletions affecting either of two cysteine-rich domains in the extracytoplasmic portion of the molecule. The clonal nature of these alterations, and the fact that identical alterations were seen in more than one tumor, suggests a role for these mutant receptor proteins in tumorigenesis. Further, these studies document the existence of tumor-specific cell surface molecules resulting from somatic mutation.

The epidermal growth factor receptor (EGFR) gene is the cellular homolog of the v-*erbB* oncogene originally identified in avian erythroblastosis viruses (1, 2). Activation of this oncogene by gene amplification has been observed in a variety of human tumors, and in particular, those of glial origin (3-6). The one consistent difference between v-*erbB* oncogenes and the EGFR gene is that the viral oncogenes encode N-truncated versions of the normal receptor; they lack most of the extracytoplasmic domain but retain the transmembrane and tyrosine kinase domains (7). This results in a protein that is unable to bind EGF but can still phosphorylate other substrates (8), and has led to speculation that the v-ErbB proteins are oncogenic because the kinase domain is unregulated and constitutively active (1).

Amplification of the EGFR gene occurs in 40% of malignant human gliomas. In many of the tumors, the amplification is accompanied by gene rearrangements (3–6). We have now examined the effect of these rearrangements on the structure of the gene product in five human glial tumors. All five tumors were found to have deletions of the EGFR gene affecting the extracytoplasmic portion of the gene product. The fact that identical alterations were seen in more than one tumor suggests a common role for these mutant receptor proteins in tumorigenesis. Further, these studies document the existence of tumor-specific cell surface molecules resulting from somatic mutation in human solid tumors. Such molecules represent potential targets for monoclonal antibodies in diagnosis and/or therapy.

## MATERIALS AND METHODS

**Source of Tumors.** The xenografts used in the present study were derived from surgical biopsies of malignant human gliomas. Southern blotting experiments with EGFR cDNA probes have shown that five of eight xenografts examined (D245MG, D256MG, D270MG, D298MG, and D317MG) contained rearranged and amplified EGFR genes and that the same rearranged fragments were detected in the original tumor biopsy specimens as in the xenografts (9).

Cloning of the EGFR Gene. A genomic library was constructed using DNA from D320MG, a glioblastoma xenograft without EGFR gene rearrangements (9). After partial *Mbo* I digestion, DNA was cloned into the *Bam*HI site of  $\lambda$  FIX (Stratagene). The ligation product was packaged with  $\lambda$  phage extracts (Stratagene) and used to infect *Escherichia coli* C600 cells. DNA from the resulting plaques was lifted with Colony/ Plaque Screen membranes (DuPont/NEN) and screened with EGFR cDNA probes (2, 10, 11). For Southern blot analysis, tumor DNA was purified, digested with *Eco*RI, and electrophoresed through a 1% agarose gel. The DNA was transferred to nylon membrane and hybridized as described (12).

**RNase Protection Assay.** <sup>32</sup>P-labeled RNA transcripts were generated *in vitro* from subclones of EGFR cDNA by T3 or T7 RNA polymerase. RNase protection analysis was performed as described (13), except that the final volume was 10  $\mu$ l, only RNase A (12.5  $\mu$ g/ml) was used, and the RNase A and proteinase K digestions were at room temperature for 30 min. The probes and their positions in the EGFR cDNA were as follows: probe I, a 910-base-pair (bp) *Sma* I–*Cla* I fragment of pE15, nucleotides (nt) 5–915 (10); probe II, a 730-bp *EcoRI–Bam*HI fragment of pE7, nt 617–1348 (11); probe III, a 970-bp *Bam*HI–*EcoRI* fragment of pE7, nt 1348–2318.

cDNA Library Construction. First-strand cDNA was prepared from RNA by using Moloney murine leukemia virus reverse transcriptase (BRL) and random hexamer primers (Pharmacia). The second strand was synthesized by the method of Gubler and Hoffman (14). The cDNA was ligated into  $\lambda$  ZAP (Stratagene) and packaged with  $\lambda$  phage extracts (Stratagene). Sequencing by the dideoxy method used a modified form of T7 DNA polymerase (United States Biochemical).

**PCR-Based Analysis of Transcripts.** To generate firststrand cDNA, 50 pmol of the 3' primer was annealed to  $1 \mu g$ of total RNA in the presence of 400  $\mu$ M dNTPs and reverse transcriptase buffer (BRL); after addition of 20 units of Moloney murine leukemia virus reverse transcriptase (BRL), the mixture was incubated at 37°C for 10 min. PCR was

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; EGFR, EGF receptor; ORF, open reading frame; nt, nucleotide(s).

<sup>&</sup>lt;sup>†</sup>Present address: Department of Microbiology and Immunology, Thomas Jefferson University, Philadelphia, PA 19107.

<sup>&</sup>lt;sup>¶</sup>To whom rep. int requests should be addressed.

carried out on this sample by adjusting the final MgCl<sub>2</sub> concentration to 2.5 mM, adding 50 pmol of the 5' primer and 2.5 units of Taq DNA polymerase (Cetus), and bringing the volume to 50  $\mu$ l. Thermal cycling was done at 93°C for 1 min, 40°C for 2 min, and 72°C for 2 min, for 35 cycles. The sample was then electrophoresed in a 2% agarose gel. The band was excised, and the DNA was eluted and subcloned into pBluescript (Stratagene) for sequencing. At least two clones from each PCR were sequenced, and some samples were sequenced directly (15). For tumors D270MG, D317MG, and D397MG, primer set A was used: 5'-AGTCGGGCTCTG-GAGGA-3' and 5'-CACTGATGGAGGTGGAGT-3'. For tumors D256MG, D298MG, and D328MG, primer set B was used: 5'-(CTG)CAGGTCTGCCATGCCTTG-3' and 5'-(GGT)ACCATCCCAGTGGCGATG-3'. The sequences in parentheses were not present in the EGFR sequence and were added to complete either a Pst I or a Kpn I restriction site.

## RESULTS

Internal Deletions in the EGFR Gene in Human Glial Tumors. To analyze the structure of the altered EGFR genes, we first constructed a map of the gene. A genomic DNA phage library was made from D320MG, a human glial tumor xenograft that has a 10-fold amplification of this gene but no detectable rearrangements. This library was screened with fragments from a 5.5-kilobase (kb) EGFR cDNA clone (2, 10, 11), and 48 phage clones were obtained and used to assemble an EcoRI map of the EGFR gene (Fig. 1A). The clones spanned the entire gene except for two small gaps. Southern blot hybridization using genomic DNA digested with various enzymes confirmed that the clones identified all EcoRI fragments within the gene (including those containing the gaps) and that the map deduced using these tumor-derived phage clones was not rearranged in comparison to normal genomic DNA. During the course of this work, a map of the EGFR gene was published by Haley et al. (16); the map in Fig. 1 agrees with theirs except that the sizes of some of the fragments differ (for unknown reasons) and our map includes several small EcoRI fragments not included previously.

Phage clones of the gene were used to probe Southern blots of *Eco*RI-digested tumor DNA (Fig. 1*C*). Clones ERG 9, 40, and 24 demonstrated no amplification in D245MG (Fig. 1*C*, lane a). However, the phage ERG 26 probe, which hybridized to seven *Eco*RI fragments in normal DNA, revealed only two amplified fragments in this tumor: the normal 8.0-kb fragment and an abnormally migrating 1.7-kb fragment. Several bands that do not appear to be amplified belong to the remaining EGFR allele, which has not undergone any amplification or rearrangement. These results suggested that the 5' portion of the EGFR gene was deleted in D245MG.

Phage clones 40 and 24 revealed deletions and rearrangements in tumors D270MG (Fig. 1C, lane c) and D317MG (data not shown) but not in D256MG or D298MG. Conversely, clones 26 (Fig. 1C) and 29 (data not shown) demonstrated a rearrangement in tumors D256MG (Fig. 1C, lane b) and D298MG (lane e) but not in tumors D270MG or D317MG. Use of EcoRI subfragments of phage clone 26 showed the 2.5-kb fragment in tumor D256 (Fig. 1C, phage clone 26, lane b) to be due to a rearrangement affecting the 8.0-kb and 1.8-kb fragments (Fig. 1A). That the normal-size 8.0-kb fragment was still present at increased copy number in tumor D256MG could be due to the presence of several copies of chromosome 7, which were detected cytogenetically in metaphase spreads from this xenograft (9). Such numerical increases in chromosome 7 are common in glial tumors (17). Alternatively, the amplification units in D256MG may be heterogeneous, with some containing the normal 8.0-kb fragment and others containing the rearranged 2.5-kb fragment. From Southern blots probed with genomic phage clones, the approximate areas of deletion within the EGFR gene in each tumor were deduced (Fig. 1B). While none of these rearrangements were identical, the deletions in D270MG and D317MG centered around a locus at the 5' end of the gene, whereas the deletions in D256MG and D298MG centered around a more distal locus. The deletion in D245MG appeared to involve most of the proximal half of the gene. Cloning experiments and the screening of a somatic cell hybrid panel confirmed that the 3' end of the EGFR gene in D245MG had fused to a new locus that was located at 7p12-14, the same chromosomal position as the EGFR gene (data not shown) (10).

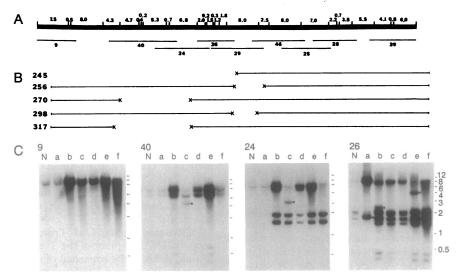
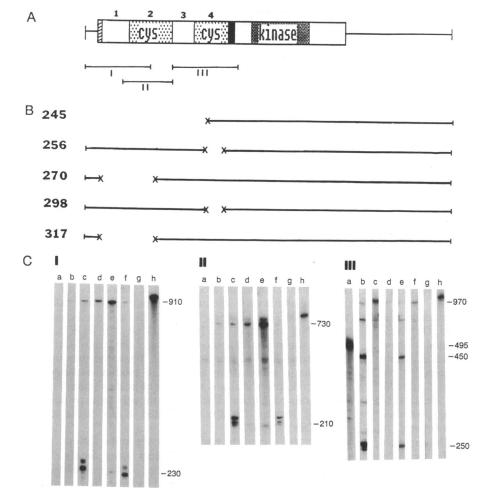


FIG. 1. Deletions in the EGFR gene in human gliomas. (A) EcoRI map of the human EGFR gene. Fragment sizes are indicated in kilobases. Representative phage clones used to assemble the map are shown below it. (B) Deduced EGFR gene structure of five glioma xenografts. Numbers at left correspond to the glial tumors described in the text. Solid lines indicate sequences present in the tumors, and the approximate points of deletion are indicated (x). (C) Southern blot hybridization with phage clones demonstrate deletions in the EGFR gene. Blots were hybridized with radiolabeled phage inserts; numbers above each blot refer to the phage clone used as hybridization probe. Rearranged fragments are indicated by asterisks. Tick marks to the right of each autoradiograph refer to sizes of marker fragments in kilobases given at extreme right. DNA was from normal cells (lanes N) or from D245MG (lanes a), D256MG (lanes b), D270MG (lanes c), D320MG (lanes d), D298MG (lanes e), or D317MG (lanes f).

Medical Sciences: Wong et al.

Analysis of the Transcripts by RNase Protection. The probes used to analyze total RNA from these tumors are shown in Fig. 2A. Probe I contained a 910-bp fragment from the 5' end of the cDNA and yielded a fragment of the same size from all RNA samples tested except for tumor D245MG (Fig. 2C). In the samples from the tumors with rearrangements, the normal-size fragments were probably a result of transcription from the remaining normal EGFR gene. In tumors D270MG (Fig. 2C, probe I, lane c) and D317MG (lane f), the most intense fragment detected was  $\approx$ 230 bp; these new fragments appeared to be expressed at a considerably higher level than the normal-size 910-bp fragment, especially when the difference in size was taken into account.

Probe II revealed fragments of 210-215 bp in tumors D270MG and D317MG, as well as a less intense 730-bp fragment corresponding to the normal transcript (Fig. 2C, probe II, lanes c and f). The cluster of closely migrating fragments (rather than a single band) was most likely due to imperfect cleavage of the duplex; RNase protection experiments in which similar RNA probes were hybridized to cloned genomic fragments also produced a cluster of fragments (data not shown). Probe III demonstrated several abnormal fragments in D256MG and D298MG, including one of 250 bp (Fig. 2C, probe III, lanes b and e). Tumor D245MG yielded a single band of 495 bp (Fig. 2C, probe III, lane a). Probes specific for the 3' half of the receptor (including the tyrosine kinase domain and autophosphorylation sites) showed no abnormalities in any of the five tumors, indicating that there were no point mutations detectable by this method (data not shown). Although probes I and II span 1343 bp of the cDNA, a total of only  $\approx$ 440 bp was detectable by this assay in tumors D270MG and D317MG. If the deletion in the transcript was continuous, this is consistent with an  $\approx$ 900-bp



deletion near the 5' end of the transcript. While the genomic deletions were not identical (Fig. 1), it appeared that the two tumors had lost the same exon sequences. Similarly, probe III spans 970 bp, but the sum of the two fragments found in tumors D256MG and D298MG was only 700 bp, suggesting a 270-bp deletion. D245MG contains a 495-bp fragment demonstrating that only this portion of the transcript was retained.

The Predicted Protein Product in D245 is Similar to v-ErbB. To determine the nature of the aberrant transcript from D245MG cells, we constructed a cDNA library using poly(A)<sup>+</sup> RNA from the D245MG xenograft. Analysis of 17 different cDNA clones selected with a probe for the EGFR gene showed that they were identical to published sequences of EGFR mRNA from nt 885 to the 3' end of the coding region (Fig. 3). The sequences 5' to nt 885 of the D245MG cDNA clone were not homologous to EGFR cDNA and were apparently derived from the upstream locus. Search of a DNA database (GenBank, release 52.0) revealed no significant homologies of this sequence to any known sequence. Open reading frame (ORF) analysis upstream of nt 885 showed numerous stop codons in all three reading frames. The single long ORF began at D245MG cDNA nt 955 (Fig. 3) and continued in the native reading frame of the EGFR. The nucleotides surrounding the first methionine codon (GC-CATGA) within this ORF were in reasonable agreement with previously described initiator sequences (18). Thus, it is likely that the translation product of this RNA has an N terminus corresponding to amino acid 543 of the EGFR and that the ORF is preceded by a long 5' untranslated sequence. Comparison of the predicted amino acid sequence from the D245MG EGFR gene with that of v-erbB oncogenes (1, 7) revealed that the protein product of the tumor is very similar

> FIG. 2. Characterization of the transcripts by RNase protection. (A) Schematic of the probes. The 5.5-kb EGFR cDNA is diagrammed above probes I-III. The hatched box is the signal peptide, followed by the extracytoplasmic domain, which is divided into four domains (nos. 1-4), with the cysteine-rich regions as indicated. The solid box is the transmembrane domain, followed by the intracytoplasmic domain, which contains the kinase domain. The thin lines are the untranslated regions. (B) Transcript structures deduced from RNase protection. Symbols are as in Fig. 1B. (C) RNase protection experiments using RNA from the glial tumor xenografts. Numerals I-III refer to the probes in A. Sizes of the protected fragments (see text) are given at right. RNA was from D245MG (lanes a), D256MG (lanes b), D270MG (lanes c), D320MG (lanes d), D298MG (lanes e), and D317MG (lanes f). Other lanes show tRNA (negative control) (lane g) and undigested probe (lane h).

2968 Medical Sciences: Wong et al.

0245         6CCARCECTETECECECCATTECCECARAGECTECCTCARAGEAGETTERCCCCCEFGETTERCCCCCCGGGGTTCCCCACC         75           0245         GEARCCCARTCACCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG			
D245       CORCEGNEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEG	D245	eccaecetcttececceccatteceecaectetcctcaeaecaegtcteeccceeteecteeacceecccca	ħ
D245       ACCTEGCCTAAGGAGGAGATTGAGTGACTCTCACTCACCTCGGTGTTGCTCTTTEAAAGTGGGCCTTGGCACC       300         D245       ACCATGACTCCCCATCCCCACCACACTCGGTGTTTTGGGTCTTCTTTEAAAGTGGCGCCTTGCCACAC       375         D245       CTTGCCAGGATGCAGATTTGAATACTATAGTGAAGTCCTGTACATGAGAAAATGATGCTTTTAGGAGGAAAAAAA       450         D245       CTTGCCAGGATGCAGATTTGAATACTATAGTGAAGTCCTGTACATGAAGAAATGATGCTTTTAGGAGGAGAAAAAAA       450         D245       CTTGCCAGGAAGAAGCCCCTTCAATGCGACATGCAGAGAATGCTGCTTTAGGAGGGAG	D245	GRAGECCAGECACCEGGGGGETCATEGGETCAGGCTGCGGGGGCCCTCGGCACCTTCTCCCGCCCCGGGGTTCCCACG	150
D245       AGCATGAACTCCCCATCCTCAGCATCCCATCAGGTGTTTTGGGTCTTCAACCTAAAATTCTATCTTACAAGATC       375         D245       CTTGCCAGGATGCAGATTTGAATACTATAGTGAAGTCTGTACATGAAGAAATGATTCTATCTTACAAGATC       375         D245       CTTGCCAGGATGCAGATTTGAATACTATAGTGAAGTCTGTACATGAAGAAATGATCTTTTAGGAGTGGGCGCC       525         D245       CCTGGAAGACAGAATTATTGTTGTGGCTATCCATTGGCTCCTGAGGCTCTAACCAGAGATGGGGGGCACCCTTTAGTA       600         D245       CCTGGGGAGTGACTGTTGCCCATAAGGTACTGGACATCGAAGCAGCCCCAGCTCCTTAAGCTGCGGG       675         D245       CCAGGGGGAGTGACTGTTGCCCATAAGGTACTGGACATCAACTTTGGCGCACATTACAACATAAGAACTGCAGGGG       675         D245       TCCTGGTGCATCGCTGCTGCAGACGACATGAAGAAGAAGCAGCCCCAGCTCCTTAAGCAGCTGCAGAGAG       750         D245       GGTGTAATCCCAGTGGAAGACTGAATGGAGGCTCAAAGAGGTTTGGCGCAACTAACATAGAAGCTGCAAGGAAGTTTTTTCT       825         D245       CAGGAGGGAAACTGAATGGAAGAATGAATGAATGAATGAGTGCTGCCCAAGGAAGTTATGCCTTCCAGGAGAGAGTT       825         D245       CAGGAGGGAAACTGAATGGAGGAACTGAACGAGCTCAACAGGCAAGGCAAGCAA	D245		225
D245       CTTGCCAGGATGCAGATTTGAATACTATAGTGAAGTCGTACATGAAGAAATGATGCTTTTAGGGAGGAAAAAAA       450         D245       AAGGTAATAACAACCTTCAAGCCCCTTCATCTCAACTCGGCATAAACAAGGCAAGAATTCTGAGAGTGGCCGCC       525         D245       CCTGGAAGCAGAATTATTCTTGTGGCTATCCATTGGCTCCTGAGGCTCTAATCAGAGATGGCGCACCCTTTAGTA       600         D245       CCAGGGGGAGACTGTTGCCCATAAGGAACTGCAACTGCAGGCTCTAATCAGAGATGGGGCACCTTTAGTA       600         D245       CCAGGGGGGAACTGTTGCCCATAAGGAACTGCAACTGCAAGCACTCCAAGCAACGACGCCCCTTAAGCAGCGCCCTTAAGAACTGCCGCTGGCGCACGAGGAACTGCAGCGAAGAATTGATCGCTTCTGAGGCAACGAACTGCAGGAGGAACTGCAGGAAGAATGGAAGCAAGAGAATGGAAGCAAGC	D245	ACCTGGCCTANGEAGGCAGATTGAGTGACTCTCACTCACCACTGGTGTTGCTCTTTGAAAGTGGCGCCTTGGCACC	300
D245       AAGGTAATAACAACCTTCAAGAGCCCCTTCATCTCAACTCGGCATAAACAAGGCAAGATTCTGAGAGTGGCCGCC       525         D245       CCTGGAAGCAGAATTATTCTTGTGGCTATCCATTGGCTCCTGAGGCTCTATCCACAGAGTGGCGCCCCTTTAGTA       600         D245       CCAGGGGGAAGTGACTGTTGCCCATAAGGTAGTGGCACACTCAATCAGAGAGCGCCAGCTCCTTTAGTA       600         D245       CCCAGGGGGAAGTGACTGTTGCCCCATAAGGTAGTGGCACACTCAATCAGAGCAGCCCCAGCTCCTTTAGTA       607         D245       TCCTGGTGCATCTGCTGCCAAAAGGTAGCAGAGCTTTGGCGACATTACAACTAAGGAAGTGCTGCTGC       675         D245       GGTGTAATCCCAGTGCAAGACTGAAGAAGAACTGAAAGGAAGTTATGCCTTCTCAGGAATGCTGTTTTTTTCT       825         D245       CAGGAGGGGAACATGCAGGAGGAAGAATGAAGAAATGAAGCAGAGCTTCGGGGAAGAATGCTGTTTTTTTT	0245	AGCATEMACTCCCCATCCACAATCCCCATCAGETGTTTTEGGTCTTCAACCTAAAATTCTATCTTACAAGATC	375
D245       CCTGEAAGCAGAAATTATTCTTGTGGCTATCCATTGGCTCCTGAGGCTCTAATCAGAGATGGGGCACCTTTAGTA       600         D245       CCAGGGGGAGTGACTGTTGCCCATAAGGTACTGGACATCGAGGCTCTAATCAGAGATGGGGCACCTTTAGTA       600         D245       CCAGGGGGACATGCTGCTGCCCATAAGGTACGGACATCGAGACCACGCCCCACTCCCTTAAGCTGCTGGC       675         D245       TCCTGGTGCATCTGCTGCCGACATGAAGGTAGGAGACACGCCCAGCTCCTTAAGCTGCAGAGA       750         D245       GGTGTAATCCCCGTGCGTGAAGCAAGGAAGCACACAGAGAGTTTGGCGACAAGGAAGTATGCCGTGCCCAGGGAGAGACAAGGAGAGTT       825         D245       CAGGAGGGGAACATGGAGGAAGAAATGAATGAATGATGGCTCACAGATGGTGAGGCTAAGGGCGAAGGGAAGTT       1832         D245       CAGGAGGGGAACATGGAAGGAAGGAAATGAATGGATGGCTGCCCCACAGGGCAAGGCCAAGGGCAAGGGAAGTT       1832         D245       TGTGGAGAACTCTGAGTGCATACAGTGCCACCCACAGGTGCCTGGCTCAGGCCAAGGGCAAGCCAAGGGAAGTT       1832         D245       TGTGGAGAACTCTGAGTGCATACAGTGCCACCCACAGGTGCCTGCGCTCAGGCCAAGGCAATGCACCTGCACAGGAAGCT       1907         D2455AA       GGGACCAGACAACTGTATCCAGTGTGCCCACTTACATTGACGGCCCCCACTGCGTCAAGG       1033         10245       GGGACCAGACAACTGTATCCAGTGTGCCCACTTACATTGACGGCCCCCACTGCGTCAAGG       1965         D2455AA       G P D H C I Q C A H Y I D G P H C Y K       1965	D245	CTTCCCACCATCCACATTTCAATACTATACTCAACTCATCA	450
D245       CCAGGGGGAGTGACTGTTGCCCATAAGGTACTGGACATCAACTTTCAAGAGCAGCCCCAGCTCCTTAAGCTGCTGG       675         D245       TCCTGGTGCATCTGCTGACTGTCAACTGTAGAAGATAGCAGAAGCTTTGGCGACATTACAACATAAGAACTGCAGAAGA       750         D245       GGTGTAATCCCCAGTGCAAGCTGAATCGAAGAAGCTCAAAAGGCAGCCCAACATTACAACATAAGAACTGCAGAAGA       750         D245       GGTGTAATCCCCAGTGCAAGAACTGAATGCAAGAAGCTCAAAAAGGCAGCTTTGGCGAACATTACAACATAAGAACTGCTAGTTTTTTCT       825         D245       CAGGAGGGGAACATGCAAGAAGGAAATGAATGAAGTGCTGACGCAAGGAAGATGATGGTGAAGATCTCAGGAAGGGAAGAATGAAT	D245		525
0245         TCCTGETECATCTGCTGACTITICATGTAGAAGATAGCAGAGCTTTGGCGACATTACAACATAAGAACTGCAGAGA         750           0245         GETGTAATCCCAGTGCAAGACACTGAATCGAGAGACTCAAAAGATAGCAGTATGCCTTCTCAGAAGACTGCTGCTTCAGGAGAGAATGCAGGAGAAATGGAAGAAATGGAAGAAGTGATGGCTGCCTCAGGAGAGATGCCAGGAGGAGAATGGAAGAAATGGAAGAAGTGCCAGGAGGAGGAGAGAACTGCAGGAGGAGAAATGGAAGAAGTGCCAGGAGGAGGAGAGAACTGCAGGAGGAGGAAGAACTGGAGGAGGAGGAAGAACTGGAGGAGGAGGAAGAACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA	0245	CCTGGAAGCAGAAATTATTCTTGTGGCTATCCATTGGCTCCTGAGGCTCTAATCAGAGATGGGGGCACCTTTAGTA	600
0245         GETGTAATCCCAGTGCAAGACTGAATCGAAGACTCAAAAAGGAAGTTATGCCTTCTCAGAATGCTGTTTTTTTT	0245	CCAGGGGAGTGACTGTTGCCCATAAGGTACTGGACATCAACTTTCAAGAGCAGCCCCAGCTCCTTAAGCTGCTGG	675
D245         CAGGAGGGGAACATGCAGGAGGAACAATGAATGATGGCTCACAGATGGTGAGATCCTCAGGCCACAGGGGGAACATGCAGGAGGGAACAATGAATG	0245	TCCTGGTGCATCTGCTGACTTTCATGTAGAAGATAGCAGAGCTTTGGCGACATTACAACATAAGAACTGCAGAGA	750
EGFR        GGACAMAGTGCAMGCTTCTGGAGGGGTGAGGCAAGGGAGGTT         1832           D245         TGTGGAGAACTCTGAGTGCATACAGTGCCAGGGCGACGAAGGCCATGAGCCAAGGGCAGTGAACACTCACCTGCACAGGAGGG         975           EGFR         TGTGGAGAACTCTGAGTGCATACAGTGCCACGAAGTGCCTGCC	0245	egtgtaatcccagtgcaagactgaatccagactcaaaaggacgtatgccttctcagaatgctgtttttct	825
EGFR      GGACAMETECAMECTTCTERESETEACCAAGEAGETT       1852         D245       TETGEAGAACTCTEAGTECATACAETECCACCAAGETECCTECCTCAGECCATEAACATCACCTECACAGEAGE       975         EGFR       TETGEAGAACTCTEAGTECATACAETECCACCAAGETECCTECCTCAGECCATEAACATCACCTECACAGEAGE       1907         D2455AA       BEACCAGACAACTETATCCAETECCCACTACATTEACCECCCCACEGETCAAGE       1033         EGFR       GEGACCAGACAACTETATCCAETECCCACTACATTEACCECCCACEGETCAAGE       1035         D2455AA       G P D N C I Q C A N Y I D G P N C Y K       1965	D245		900
EGFR         TETGEAGAACTCTGAGTGCATACAGTGCCACAGAGTGCCTGCCTCAGACCATCACCTGCACAGAGAG         1907           D2455AA         N         N         N         I         T         C         T         6         R           D2455AA         GGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAGA         1033         1033           EGFR         GGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAGA         1035           EGFR         GGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAGA         1035           EGFR         GGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAGA         1035           EGFR         GGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAGA         1965           D2455AA         G         P         N         C         N         Y         1         D         G         N         1965	EGFR		1832
EGR         TETERARAACTCTRAGTGCATACAGTGCCACCAAAGTGCCTGCCTCAGGCCACACACA	0245	TETERAGAACTCTEAETECATACAETECCACCCAEAETECCTECCTCAEBECCA <u>TB</u> AACATCACCTECACAEBACG	975
EGFR GEGACCAGACAGCTGTATCCAGTGTGCCCCCTACATTGACGGCCCCCACTGCGTCAGG 1965 D245AA G P D N C I Q C A N Y I D G P N C V K		TETECAGAACTCTEAETECATACAETECCACCCACAETECCTCCCCCACAEAACATCACCTECACAECACE	1907
EGFR GEGACCAGACAGETGTATCCAGTGTGCCCCACTACATCGAGGCCCCCCACTGCGTCAGG 1965 D245AA G P D N C I Q C A N Y I D G P N C V K	0245	CORACCARACAMCTGTATCCAGTGTGCCCACTACATTGACGGCCCCCACTGCGTCAAG	1033
		CORACCARACAACTETATCCARTETECCCACTACATTEACECCCCCCACTECETCAAE	1965

FIG. 3. cDNA sequence of the EGFR-related mRNA from D245MG. D245, nucleotide sequence of a clone derived from tumor D245MG mRNA; EGFR, normal EGFR cDNA sequence (2); D245AA, deduced amino acid sequence from D245, which is aligned with that of the v-ErbB oncoprotein [a fusion protein between viral and EGFR sequences (1); only the EGFR-related sequences are shown]. The presumed initiator codon for tumor D245MG is underlined. Numbers at right refer to the D245 sequence.

(and in fact only 8 amino acids longer) than several avian retroviral gene products described previously.

In-Frame Deletions Occur in the Extracytoplasmic Domain of the Other Four Tumors. We used the PCR to generate cDNA fragments that encompassed the abnormalities in the other tumors. The RNase protection data were used to guide the choice of primers. When RNA from tumors D270MG and D317MG was analyzed with one set of primers, a major fragment of 240 bp was produced (Fig. 4A, lanes a and b), while the normal-size 1100-bp fragment was seen only faintly (data not shown), a pattern consistent with the size of the deletions inferred from RNase protection. Another glial tumor xenograft, D397MG, had a Southern blot pattern very similar to that of D270MG when probed with the EGFR cDNA. It was analyzed in this PCR experiment and produced a pattern similar to that from D270MG and D317MG (Fig. 4A, lane c). A second set of primers revealed a fragment of 220 bp in tumors D256MG and D298MG (Fig. 4A, lanes d and e, respectively), instead of the normal 440-bp fragment that was found with D320MG RNA (lane f); the 220-bp fragment size was consistent with the RNase protection data.

PCR products from these tumors were subcloned and sequenced. The sequences from tumors D270MG, D317MG, and D397MG are identical (Fig. 4B). The 5' portion of these clones matches the published EGFR cDNA sequence (2) until nt 274, which corresponds to the end of the first exon (19), but then continues with sequence normally present at nt 1076. The sequence derived from tumor D256MG is identical to that from D298MG; the sequence matches the published EGFR sequence upstream of nt 1817, which corresponds to the end of an exon, but then continues with sequence normally present at nt 2067 (Fig. 4C). In all cases, the deletions did not alter the reading frame and reconstituted a glycine codon at the deletion site (Fig. 4 B and C).

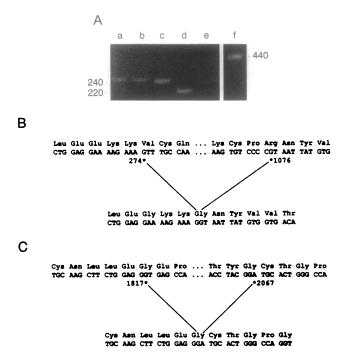


FIG. 4. PCR analysis of the gene products from the altered region. (A) Gel electrophoresis of PCR products. The gel shows 1/10th of the reaction product after 35 cycles. Lanes: a, D270MG; b, D317MG; c, D397; d, D256MG; e, D298MG; f, D320MG. Primer set A was used for D270MG, D317MG, and D397; primer set B was used for D256MG, D298MG, and D320MG. Numbers refer to the sizes of fragments in base pairs as judged from coelectrophoresed markers. (B) Sequence of the PCR products from tumors D270MG, D317MG, and D397. Numbers and asterisks in the upper line refer to the EGFR cDNA nucleotides (2) flanking the deleted area. The sequences to the right and left of those shown were the same as in normal EGFR cDNA. (C) Sequence of the PCR products from tumors D256MG and D298MG.

## DISCUSSION

Previously, we examined the size of the proteins produced in these xenografts and their effect on EGF binding affinity (9). As would be predicted from the sequence data presented here, the protein from D245MG was unable to bind EGF but could be autophosphorylated. The major polypeptide detected with antibodies against the intracytoplasmic domain of the receptor was 77 kDa after deglycosylation, consistent with the 72 kDa predicted from the cDNA. The proteins immunoprecipitated from tumors D270MG and D317MG were  $\approx 145$  kDa, while the predicted polypeptide was 100 kDa. The difference in size from that predicted from the cDNA sequences may reflect the degree of glycosylation, as there is a similar discrepancy between the predicted molecular mass of normal EGFR and its mobility in SDS/ polyacrylamide gels. The proteins found in tumors D256MG and D298MG are 120 kDa, similar to the size predicted from the cDNA sequence. The binding affinity for EGF in these tumors was more than one-third that of the control cell lines with normal EGFR gene sequences. Thus, the deletions do not seem to substantially affect EGF binding, although the normal receptors that are still present in these tumors complicate this conclusion. The domain thought to be responsible for EGF binding (20) was not deleted in any of the tumors except for D245MG.

Our results confirm and extend the observations previously made by us and other workers using Southern blot assays on glial tumors. Libermann et al. (3) noted that two glial tumors with DNA amplification had apparent rearrangements that were detected with cDNA probes from the extracytoplasmic domain. Similarly, Yamazaki et al. (5) and Malden et al. (6) detected rearrangements in two glial tumors and found that these produced protein products of abnormally small size. Two laboratories have also described the deletion involving nt 275-1075. Yamazaki et al. (21) cloned the cDNA corresponding to the EGFR from one glial tumor, found that this deletion was present, and verified that the gene itself was also altered. Sugawa et al. (22) used a PCR-based approach to detect this deletion in six glial tumors. Thus, this alteration has been described in 10 tumors so far.

One of the unexpected observations made here is that identical deletions (at the gene product level) occurred in tumors arising in different patients. Previous studies with these tumor xenografts have shown that these proteins are expressed on the cell surface (9). Now we know that there is an amino acid sequence at the deletion junction of these proteins that is not present on normal cells. This implies that a truly tumor-specific cell surface molecule is present in these tumors, against which monoclonal antibodies might be generated. Indeed, we have shown (23) that antibodies raised against a synthetic peptide corresponding to the novel junction created by the deletion found in tumors D270MG or D317MG (Fig. 4B) will specifically recognize the mutant and not normal EGFR protein. Hence, it is possible to derive antibody conjugates that will be useful for tumor imaging or therapeutic purposes. The only other analogous situation in human tumors is that associated with the unique T-cell receptors and immunoglobulins present on lymphoid leukemias and lymphomas. In some cases, it has been shown that immunotherapy directed against these unique cell surface molecules can be successful (e.g., ref. 24). The mutant cell surface molecules on glial tumors present a crucial potential advantage in that the mutant molecules themselves probably play an important role in the neoplastic process.

N-terminal truncations in v-ErbB proteins are known to be important for their effect; in the tumors studied here, the alterations were confined to the N-terminal domain. In one tumor, the deletion resulted in a truncation remarkably similar to that found in v-erbB oncogenes, whose products are thought to be oncogenic because the truncation results in a protein whose kinase activity is constitutively active and unregulated. Overexpression of a protein lacking the EGFbinding domain can transform Rat-1 cells and activate the tyrosine kinase (25). The deletions in the other tumors studied occurred within one of two cysteine-rich domains in the receptor. Therefore, these alterations could result in a conformational change that might leave the receptor in an abnormally active or unregulated state. Indeed, Yamazaki et al. (21) have introduced a cDNA corresponding to the mutation seen in tumors D270MG and D317MG and found that it does have ligand-independent transforming activity.

In an earlier publication, we showed that amplification of the EGFR gene occurred in 40% of glioblastoma multiforme tumors and that increased levels of EGFR mRNA were invariably associated with gene amplification (4). With respect to EGFR expression, there appear to be two classes of tumors: those that contain amplification and those with a normal copy number of this gene. One possible explanation for this difference is that a certain subset of glial tumors are dependent on the overexpression of EGFR. Evidence for this possibility comes from experiments in which transformation of NIH 3T3 mouse fibroblasts through overexpression of the EGFR required the presence of EGF (26-28) and the demonstration that some glial cell lines contain transforming growth factor  $\alpha$  transcripts (29). In vivo, tumors with a high density of normal receptors might be activated by otherwise limiting quantities of EGF or transforming growth factor  $\alpha$ present in the tumor milieu derived from exogenous sources or through an autocrine loop. A different subset of tumors might have evolved receptors whose structure results in a molecule that is constantly active and is only partially or not at all regulated by ligand. Such independence from the normal signals controlling cellular growth is the essence of tumorigenesis.

This work was supported by the Preuss Foundation and grants from the National Cancer Institute.

- 1. Downward, J., Yarden, Y., Mayes, G., Scrace, E., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. & Waterfield, M. D. (1984) Nature (London) 307, 521-527.
- Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Down-ward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D. & Seeburg,
- P. H. (1984) Nature (London) **309**, 418-425. 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. Wong, A. J., Bigner, S. H., Bigner, D. D., Kinzler, K. W., Hamilton,
- 4. . R. & Vogelstein, B. (1987) Proc. Natl. Acad. Sci. USA 84, 6899-6903.
- 5.
- X. & Vogelstein, B. (1967) Froc. Natl. Acad. Sci. USA 44, 697–695.
   Yamazaki, H., Fukui, Y., Ueyama, Y., Tamaoki, N., Kawamoto, T., Taniguchi, S. & Shibuya, M. (1988) Mol. Cell. Biol. 8, 1816–1820.
   Malden, L. T., Novak, U., Kaye, A. H. & Burgess, A. W. (1988) Cancer Res. 48, 2711–2714. 6.
- Gammett, D. C., Tracy, S. E. & Robinson, H. L. (1986) Proc. Natl. Acad. Sci. USA 83, 6053-6057. 7.
- 8. Kris, R. M., Lax, I., Gullick, W., Waterfield, M. D., Ullrich, A., Fridkin, M. & Schlessinger, J. (1985) Cell 40, 619-625.
- Humphrey, P. A., Wong, A. J., Vogelstein, B., Friedman, H. S., Werner, M. H., Bigner, D. D. & Bigner, S. H. (1988) Cancer Res. 48, 9 2231-2238
- Merlino, G. T., Ishii, S., Whang-Peng, J., Knutsen, T., Xu, Y. H., Clark,
  A. J., Stratton, R. H., Wilson, R. K., Ma, D. P., Roe, B. A., Hunts,
  J. H., Shimizu, N. & Pastan, I. (1985) *Mol. Cell. Biol.* 5, 1722–1734.
  Xu, Y. H., Ishii, S., Clark, A. J., Sullivan, M., Wilson, R. K., Ma, D. P., 10.
- 11. Roe, B. A., Merlino, G. T. & Pastan, I. (1984) Nature (London) 309, 806-810.
- 12. Vogelstein, B., Fearon, E. R., Hamilton, S. R., Preisinger, A. C. Willard, H. F., Michelson, A. M., Riggs, A. D. & Orkin, S. H. (1987) Cancer Res. 47, 4806-4813.
- Winter, E., Yamamoto, F., Almoguera, C. & Perucho, M. (1985) Proc. Natl. Acad. Sci. USA 82, 7575-7579. 13.
- Gubler, U. & Hoffman, B. J. (1983) Gene 25, 263-269. 14.
- Winship, P. R. (1989) Nucleic Acids Res. 17, 1266. 15.
- 16. Haley, J., Whittle, N., Bennett, P., Kinchington, D., Ullrich, A. & Waterfield, M. (1987) Oncogene Res. 1, 375-396.
- 17. Bigner, S. H., Mark, J., Bullard, D. E., Mahaley, M. S., Jr., & Bigner, D. D. (1986) Cancer Genet. Cytogenet. 22, 121-135
- 18. Kozak, M. (1987) Nucleic Acids Res. 15, 8125-8148.
- Ishii, S., Xu, Y. H., Stratton, R. A., Roe, B. A., Merlino, G. T. & Pastan, I. (1985) Proc. Natl. Acad. Sci. USA 82, 4920–4924. 19.
- Lax, I., Burgess, W. H., Bellot, F., Ullrich, A., Schlessinger, J. & Givol, D. (1988) *Mol. Cell. Biol.* 8, 1831–1834. 20.
- Yamazaki, H., Ohba, Y., Tamaoki, N. & Shibuya, M. (1990) Jpn. J. 21. Cancer Res. 81, 773-779.
- 22. Sugawa, N., Ekstrand, A. J., James, C. D. & Collins, V. P. (1990) Proc. Natl. Acad. Sci. USA 87, 8602–8606.
- Humphrey, P. A., Wong, A. J., Vogelstein, B., Zalutsky, M. R., Fuller, 23. G. N., Archer, G. E., Friedman, H. S., Kwatra, M. M., Bigner, S. H. & Bigner, D. D. (1990) Proc. Natl. Acad. Sci. USA 87, 4207-4211.
- Janson, C. H., Tehrani, M. J., Mellstedt, H. & Wigzell, H. (1989) Cancer 24. Immunol. Immunother. 28, 222-232.
- Haley, J. D., Hsuan, J. J. & Waterfield, M. D. (1989) Oncogene 4, 25. 273-283.
- Di Fiore, P. P., Pierce, J. H., Fleming, T. P., Hazan, R., Ullrich, A., 26. King, C. R., Schlessinger, J. & Aaronson, S. A. (1987) Cell 51, 1063-1070
- Velu, T. J., Beguinot, L., Vass, W. C., Willingham, M. C., Merlino, G. T., Pastan, I. & Lowy, D. R. (1987) *Science* 238, 1408–1410. 27.
- Riedel, H., Massoglia, S., Schlessinger, J. & Ullrich, A. (1988) Proc. Natl. Acad. Sci. USA 85, 1477–1481. 28.
- 29. Nister, M., Libermann, T. A., Betsholtz, C., Pettersson, M., Claesson-Welsh, L., Heldin, C.-H., Schlessinger, J. & Westermark, B. (1988) Cancer Res. 48, 3910-3918.