4050–4056 *Nucleic Acids Research, 1996, Vol. 24, No. 20*

A 1.8 kb alternative transcript from the human epidermal growth factor receptor gene encodes a truncated form of the receptor

Jill L. Reiter and Nita J. Maihle*

Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, Minnesota 55905, USA

Received May 30, 1996; Revised and Accepted August 16, 1996

DDBJ/EMBL/GenBank accession nos U48722–U48729

ABSTRACT

The epidermal growth factor receptor (EGFR) is encoded by the c-*erbB*1 proto-oncogene and plays an important role in the control of cell growth and differentiation. To study the potential growth regulatory role of soluble EGF receptors, we have isolated cDNA clones encoding a truncated, secreted form of the human EGFR. The 5′ sequence of this cDNA is identical to the EGFR transcript encoding the full-length receptor through exon 10. The unique 3′ sequence encodes two additional amino acid residues before encountering an in-frame stop codon, a poly(A) addition site and a poly(A)$^+$ tail. Sequence comparison with genomic DNA sequences demonstrates that this alternative transcript arises by read-through of a splice donor site. As a result, this transcript encodes a portion of the extracellular ligand-binding domain, but lacks the transmembrane domain and the intracellular tyrosine kinase catalytic domain present in the EGFR. Conditioned medium from transfected fibroblast cells contains a 60 kDa protein that is specifically immunoprecipitated by an EGFR monoclonal antibody. These findings demonstrate that alternative processing of the human EGFR transcript produces a secreted product composed of only the extracellular ligand-binding domain.

INTRODUCTION

The epidermal growth factor receptor (EGFR) plays an important role in the control of cell growth and differentiation. Understanding the function of this receptor in tumorigenesis is of great interest because the overexpression of the EGFR in human carcinomas is frequently associated with a poor prognosis. The EGFR is encoded by the c-*erb*B1 proto-oncogene (1,2) and is structurally related to three receptor tyrosine kinases, known as ErbB2/Neu (3), ErbB3 (4,5) and ErbB4 (6). These receptors are encoded by distinct genes and together, make up the c-*erb*B family of proto-oncogenes.

The EGFR includes three functional domains: an extracellular ligand-binding domain, a transmembrane domain, and a cytoplasmic tyrosine kinase domain. The extracellular domain can be further divided into four subdomains (I–IV), including two cysteine-rich regions (II and IV) and two regions (I and III) involved in ligand-binding (7,8).

The 170 kDa human EGFR is encoded by two major transcripts of 5.8 and 10.5 kb (2). Additional alternatively spliced transcripts of approximately 2.6–2.7 kb have been identified in normal chicken and rat tissues; these transcripts encode secreted, truncated receptors containing only the extracellular ligand-binding domain (9,10). Furthermore, soluble EGF receptors occasionally arise from aberrant transcripts, as exemplified by the epidermoid carcinoma cell line, A431 (2). In A431 cells, the EGFR gene is amplified and rearranged, and a 2.8 kb transcript results from a translocation between the 5′ region of the EGFR gene and an unidentified region of genomic DNA (2,11,12).

Soluble truncated receptors lacking their transmembrane and cytoplasmic domains have also been reported for ErbB2 and ErbB3 (13,14). Moreover, many transmembrane growth factor and cytokine receptors have been reported to have analogous soluble, ligand-binding receptor forms detectable in the culture supernates of cell lines, and in biological fluids such as serum and urine (15). The widespread distribution of soluble receptors suggests that these molecules may have important physiological roles.

Our laboratory was involved in the initial discovery and characterization of the soluble truncated form of avian ErbB1 (9), which was subsequently demonstrated to have growth inhibitory potential *in vitro* (16). To study the potential growth regulatory role of soluble EGF receptors in human carcinomas, we have isolated cDNA clones encoding a truncated, secreted form of the human EGF receptor.

MATERIALS AND METHODS

Reagents

The EGFR cDNA clone, pXER, was provided by G. Gill (17,18). Monoclonal antibodies which specifically recognize the extracellular domain of the EGFR were as follows: R1 (Amersham RPN.513), LA1 and LA22 (Upstate Biotechnology Inc. 05-101 and 05-104), and 528 and 225 (Oncogene Sciences, Ab-1 and Ab-5).

DNA probes

A 1.9 kb long cDNA probe corresponding to the ligand-binding domain (LBD) of the EGFR was synthesized by the polymerase chain reaction (PCR) from pXER. The forward primer was: 5′-TCGGGGAGCAGCGATGCGAC-3′, corresponding to bp 174–193. The reverse primer had the sequence 5′-CCATTCGTTGGACAGCCTTC-3′ representing bp 1986–2105. Nucleotide numbering is according to Ullrich *et al.* (2) unless stated otherwise.

* To whom correspondence should be addressed

Amplification was performed for 35 cycles (94°C 1 min, 65°C 1 min, 72°C 3 min) with a final extension at 72°C for 10 min. A 768 bp *Eco*RI fragment from pXER was gel purified and used as the intracellular kinase domain (KD) probe. The LBD and KD probes were radiolabeled with [α-^{32}P]dCTP using a random primer DNA labeling kit (Gibco BRL) according to the manufacturer's instructions.

cDNA library screening and analysis

A human placenta cDNA library (Clontech, catalog # HL1144x) was screened for clones encoding only the ligand-binding domain of the EGFR. Duplicate nitrocellulose filters containing 640 000 recombinant phage were hybridized separately with radiolabeled LBD or KD probes. The hybridizations were performed in a solution containing 6× SSC, 5× Denhardt's, 7.5% dextran sulfate, 0.5% N-lauryl sarcosine, and 100 µg/ml salmon sperm DNA at 65°C. Filters were washed in 0.1× SSC and 0.1% N-lauryl sarcosine at 65°C and were exposed 24–72 h at –80°C with an intensifying screen. Plasmid DNA was released from the pλDR2 vector by site-specific recombination using the CRE–lox system (19). Clones were sequenced on both strands using the Taq DyeDeoxy cycle sequencing kit and the Applied Biosystems model 373A automated DNA sequencer. In GC-rich regions of the templates, 1 µl dimethyl sulfoxide was added to the sequencing reactions.

PCR analysis of intron sequences

Intron 10 of the EGFR was amplified by PCR from human genomic DNA. The forward primer was EX10F: 5′-TGACTCC-TTCACACATACTC-3′, corresponding to bp 1320–1339 in exon 10. The reverse primer had the sequence, EX11R: 5′-TTCTCAA-AGGCATGGAGGTC-3′, representing bp 1432–1451 in exon 11. Human DNA (50 ng) was mixed with 20 pmol of each primer, 100 µM of each deoxynucleotide, 2.5 U *Taq* polymerase (Boehringer Mannheim Biochemicals), and 5 µl of 10 × buffer (supplied with the *Taq*) in a total volume of 50 µl. Amplification was performed for 35 cycles (94°C 1 min, 65°C 1 min, 72°C 2 min) with a final extension at 72°C for 10 min. The PCR product was ligated into a TA-cloning vector (Invitrogen). Plasmid DNA was isolated from two independent colonies and was sequenced as above. In addition, PCR primers were designed from the flanking exon and divergent sequences in clones 281, 721, 713, 711, and 152. The oligonucleotide sequences are available upon request.

Amplification of reverse-transcribed RNA by PCR (RT–PCR)

Total cellular RNA was isolated from the human placental cell line, 3A-Sub-E (ATCC#: CRL 1584), by the guanidine isothiocyanate extraction procedure (20). Prior to the reverse transcription reaction, the RNA was treated with RNase-free DNase I (Boehringer Mannheim Biochemicals) and extracted twice with an equal mixture of phenol and chloroform. RNA (1 µg) was heat denatured at 90°C for 5 min and then reverse transcribed in a 20 µl reaction mixture (1× AMV reaction buffer, 1 mM each dNTP, 10 mM dithiothreitol, 20 U RNasin, 10 U AMV reverse transcriptase) using 0.1 µg of oligo-dT primer for 1 cycle of 24°C 10 min, 42°C 50 min, 99°C 5 min, and 4°C 5 min. *Taq* DNA polymerase and 10 pmol of the forward primer EX10F and of either reverse primer, EX11R or P161R (5′-CCAAGGGAACAGGAAATATG-3′), were added to final volume of 100 µl. Amplification was performed as described above. Products were analyzed after electrophoresis on a 5% polyacrylamide gel and staining with ethidium bromide.

Transient transfection of QT6 cells

The quail fibroblast cell line, QT6 (21), was maintained in Dulbecco's modified Eagle's medium (DMEM, Biowhittaker) containing 4.5 g of glucose per liter and supplemented with 5% fetal calf serum (FCS) and 1% chick serum. Cells were transfected transiently with 15 µg of the expression vector pDR2 containing cDNA 161 (pDR161) by the calcium phosphate precipitation technique as described previously (22).

Metabolic labeling and immunoprecipitation of ^{35}S-labeled ErbB1-S

Transfected cells from two 100 mm plates were pooled and replated in 6-well plates approximately 48 h post-transfection. The following day, cells were rinsed once in phosphate buffered saline (PBS) and labeled in methionine-free DMEM supplemented with 5% dialyzed FCS and 150 µCi/ml of [^{35}S]methionine (Promix, Amersham) at 37°C for 12 h. Conditioned medium from labeled cells was collected and centrifuged briefly to remove loose cells and debris and phenylmethylsulfonyl fluoride (PMSF) and aprotinin were added to a final concentration of 1 mM and 50 µg/ml. Cell monolayers were lysed and immunoprecipitated with the addition of 1–5 µg of monoclonal antibody as described previously (23). Samples were resuspended in 2× Laemmli sample buffer (125 mM Tris–HCl, pH 6.8, 4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 2 mM EDTA, 0.04% bromophenol blue), boiled for 5 min and separated by 10% SDS–PAGE. Gels were stained with Coomassie blue, treated with EnHance (Dupont) and dried before an overnight exposure to film.

RESULTS

Isolation of the soluble human c-*erbB*1 cDNA

We and others have observed a 1.8 kb transcript in human placental RNA that hybridizes exclusively to an EGFR extracellular domain probe (2,24, and data not shown). These results suggested that alternative transcripts encoding only the extracellular ligand-binding domain of the human EGFR might exist. Therefore, we used differential hybridization to screen an oligo-dT primed human placental cDNA library for clones that were positive for a ligand-binding domain (LBD) specific probe, but negative for a kinase domain (KD) probe. Eleven clones hybridizing exclusively to the LBD probe were purified from approximately 6×10^5 plaques (Fig. 1).

Characterization of cDNA clones

Five clones contained sequences identical to the EGFR coding region through exon 10 (Fig. 1). Sequence analysis revealed that clone 161 contained a 1593 bp insert that contained 244 bp of 5′ untranslated region and a reading frame that codes for 381 amino acids. Clones 161 and 763 were nearly identical in sequence except that clone 763 contained one additional nucleotide (C) at the 5′ end and had a much longer poly(A)+ tail, apparently added 13 bp upstream of the cleavage site in clone 161. The unique sequence at the 3′ ends of clones 161, 763, 801, 681, and 281 exhibited an in-frame termination codon (TGA) and an AATAAA sequence, followed by a poly(A)+ tail. The divergent region begins with the

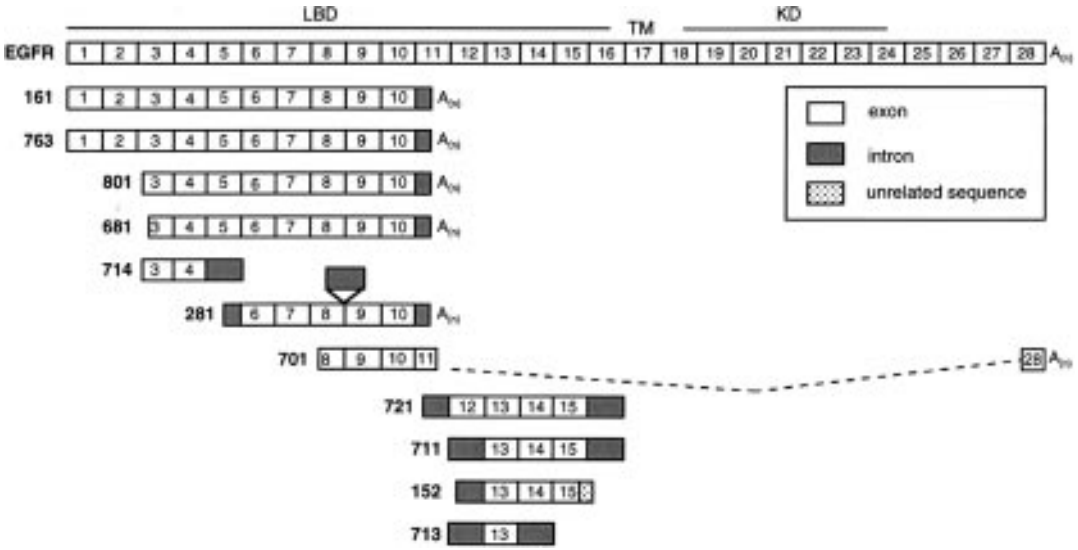

Figure 1. Schematic of overlapping cDNA clones. Exons are represented by open boxes, the shaded boxes denote intron sequences and the stippled box represents an unrelated sequence. Exon boundaries were extrapolated from Callaghan *et al.* (30). The lines above the EGFR cDNA represent the ligand-binding domain (LBD) and kinase domain (KD) specific probes used to screen the cDNA library. TM, transmembrane domain. Clones 161, 763, 801, 681, 281, and 701 contain a polyadenylation signal and a poly(A)+ tail.

sequence GTTTG, which contains the highly conserved GT dinucleotide and a G at the +5 position in the consensus splice donor site. Thus, it appears that the truncated transcript fails to splice at the 3′ end of exon 10 and reads through the intron for 117 bp to an alternative poly(A) addition site.

Six clones contained sequence insertions at various locations (Fig. 1). Sequence analysis of these clones revealed the presence of consensus splice donor or acceptor sites located between the unique sequence and sequence identical to the EGFR cDNA (Table 1). Comparison of these putative intron sequences with GenBank revealed no similarity to any known sequences. In contrast, the unique 3′ sequence in clone 152 did not diverge at a predicted exon–intron junction. Sequence comparison of this peculiar 40 bp with GenBank sequences revealed complete identity to the chorionic somatomammotropin hormone (CSH-1) gene, which is expressed at very high levels in placenta (25). Similarly, clone 701 apparently resulted from a rearrangement between the sequence 5′-CCTTTGAG-3′ located at nucleotides 1442–1449 in the extracellular region and at 4062–4069 in the 3′ untranslated region of the 5.8 kb EGFR cDNA (2), deleting the intervening region.

Clones 711 and 721 were apparently oligo-dT primed from a poly(A) sequence associated with an *Alu* repetitive element in intron 15, while clones 713 and 714 also contain a stretch of 15 As at their 3′ ends. Clones 714, 721, 711, and 713 do not contain an AATAAA sequence upstream of their poly(A)$^+$ sequences, suggesting that they were oligo-dT primed from internal poly(A) tracts present in the intervening sequences. In addition, none of these clones contained open reading frames that extended into the intron sequences. We conclude that these clones are simply artifacts of the cloning process and were derived presumably from incompletely processed transcripts that were reverse-transcribed during the construction of the cDNA library. Therefore, they do not represent additional alternatively processed transcripts.

Table 1. Human EGFR exon–intron splice junctions

Exon	5′ splice donor	Intron	3′ splice acceptor	Exon
	Ser^{161}			
4	CAGCT **gt**aagtgtcg	4		
			Thr^{168}	
		5	tctttttc**ag** TGACCAAA	6
	Lys^{311}		Cys^{313}	
8	CAAAG **gt**aggaagcc	8	atctctgc**ag** TGTGTAAC	9
	Thr^{376}		Phe^{388}	
10	CACAG **gt**ttgagctg	10	tgccttac**ag** GGTTTTTG	11
			Gly^{410}	
		11	tccaatgt**ag** TGGTCAGT	12
			Ala^{477}	
		12	gtctccgc**ag** AGGCCACA	13
	Glu^{519}			
13	GAGGG **gt**aggaggtt	13		
	Tyr^{602}			
15	TACGG **gt**gagtggaa	15		

Uppercase letters represent exon sequence, and lowercase letters represent intron sequence. The 5′-gt and 3′-ag ends of introns are in boldface. The amino acid is shown above the corresponding codon and its number in the human EGFR protein is indicated by superscript.

Comparison of human EGFR cDNA sequences

The sequence of cDNA 161 corresponding to the 5′ untranslated region and exons 1–10 was compared with published sequences of the EGFR cDNA. Overall, the sequences were identical. However, nucleotide differences were observed at 15 locations, 14 of which occurred in the G-C rich, 5′ untranslated region (Fig. 2). It is unclear if any of these base pair differences might affect promoter function. They do not appear to disrupt transcription factor Sp1 binding sites nor are they located near mapped transcription start sites (26–28). One nucleotide difference

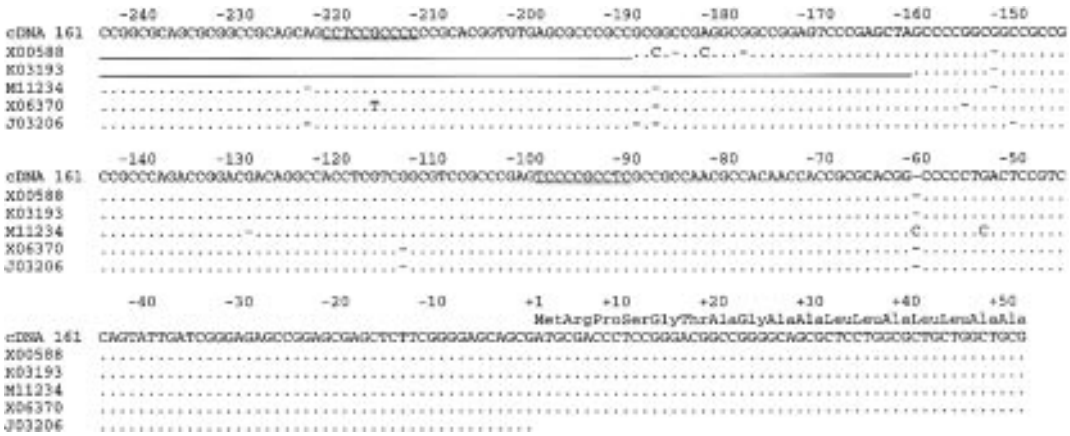

Figure 2. Multiple alignment of the 5′ region of the human EGFR gene. An alignment of the 5′ untranslated region of cDNA 161 and published EGFR gene sequences is shown. Numbers represent positions relative to the 'A' in the translation initiation codon. Gaps are represented by dashes and nucleotides identical to cDNA 161 are shown as dots. Sp1 binding sites identified by DNase I footprinting (26) are underlined. The horizontal lines indicate that no sequence information was available for that region. Sequences are labeled according to their nucleotide accession numbers and are from the following references: X00588 (2); K03193 (11); M11234 (27); X06370 (28); J03206 (43).

unique to cDNA 161 occurred in codon 134, a C-T change in the third position, which did not change the encoded amino acid (data not shown). These sequence variations may have resulted from cDNA cloning artifacts, difficulties in sequencing G-C rich regions, or they may represent sequence polymorphisms.

Exon–intron analysis

To determine whether the divergent cDNA sequences were contiguous in the genome with the flanking exons, human DNA was amplified by PCR using primers specific for the flanking exon and the unique sequences. Primers specific for exon 10 (EX10F) and the 3′ sequence of the truncated EGFR transcript (P161R) amplified a 159 bp product in both human DNA and cDNA 161 (data not shown). These results were consistent with the read-through of a 5′ splice donor site as the source of the unique sequence. To further confirm that the 3′ sequence of cDNA 161 was derived from intron 10, we amplified genomic DNA with primers specific for exons 10 and 11. The 962 bp PCR product was cloned and sequenced. Comparison of intron 10 and cDNA 161 sequences showed complete homology until the poly(A) addition site, verifying the read-through of the 5′ splice site as the origin of the novel sequence (Fig. 3).

PCR analysis with primers corresponding to exons 8 and 9 also revealed the same size product in genomic DNA compared with clone 281. However, a smaller product was obtained when an EGFR cDNA was used as the template (data not shown), indicating that clone 281 contains intron 8. Identical size products were also amplified in human genomic DNA compared with clones 721, 711, and 713 (data not shown), suggesting that these unique sequences are also intronic. The exon–intron splice junctions present in these clones are listed in Table 1 and these novel intron sequences have been deposited in GenBank (29). All of the 5′ and 3′ splice junctions reported here conform to recognized consensus sequences, and are also consistent with those reported for the chicken EGFR (30).

RT–PCR analysis of EGFR transcripts

Our initial attempts to use a 3′ specific probe from cDNA 161 did not reveal a detectable signal on northern blot analysis because of the relatively low abundance of the 1.8 kb transcript, as well as the difficulty in generating a small probe (~70 bp) with high specific activity. To determine if the transcript represented by cDNA 161 is expressed in human placenta, we developed an RNA-based PCR assay. RNA from a human placental cell line was reverse transcribed with an oligo-dT primer, and the first strand cDNA was amplified using primers specific for EGFR exon 10 and 11 or with primers specific for exon 10 and the 3′ unique sequence of cDNA 161. Specific products of the predicted size (132 bp and 159 bp respectively) were obtained, while no products were observed when the reverse transcriptase was omitted (Fig. 3C).

Soluble human c-*erbB*1 gene product

The amino acid sequence deduced from cDNA 161 predicted a 381 amino acid protein with a calculated molecular mass of 44 661 daltons. The first 24 amino acids code for a signal peptide; following cleavage by signal peptidases, the predicted molecular weight of this protein would be 42 396. The sequence encodes subdomains 1, 2 and a portion of subdomain 3 of the extracellular ligand-binding domain of the EGFR, and also retains six of 12 potential N-linked glycosylation sites (Asn, X, Ser/Thr). The final two residues, leucine and serine, are unique to this molecule and are followed by an in-frame termination codon (TGA), nine nucleotides downstream of the point of divergence with the EGFR cDNA. As a result, the predicted product lacks the transmembrane domain and the intracellular tyrosine kinase catalytic domain present in the EGFR. We have named this product soluble ErbB1 (ErbB1-S) since it is structurally related to the avian c-*erb*B1 soluble product and it is not yet known if this truncated receptor is able to bind to EGF.

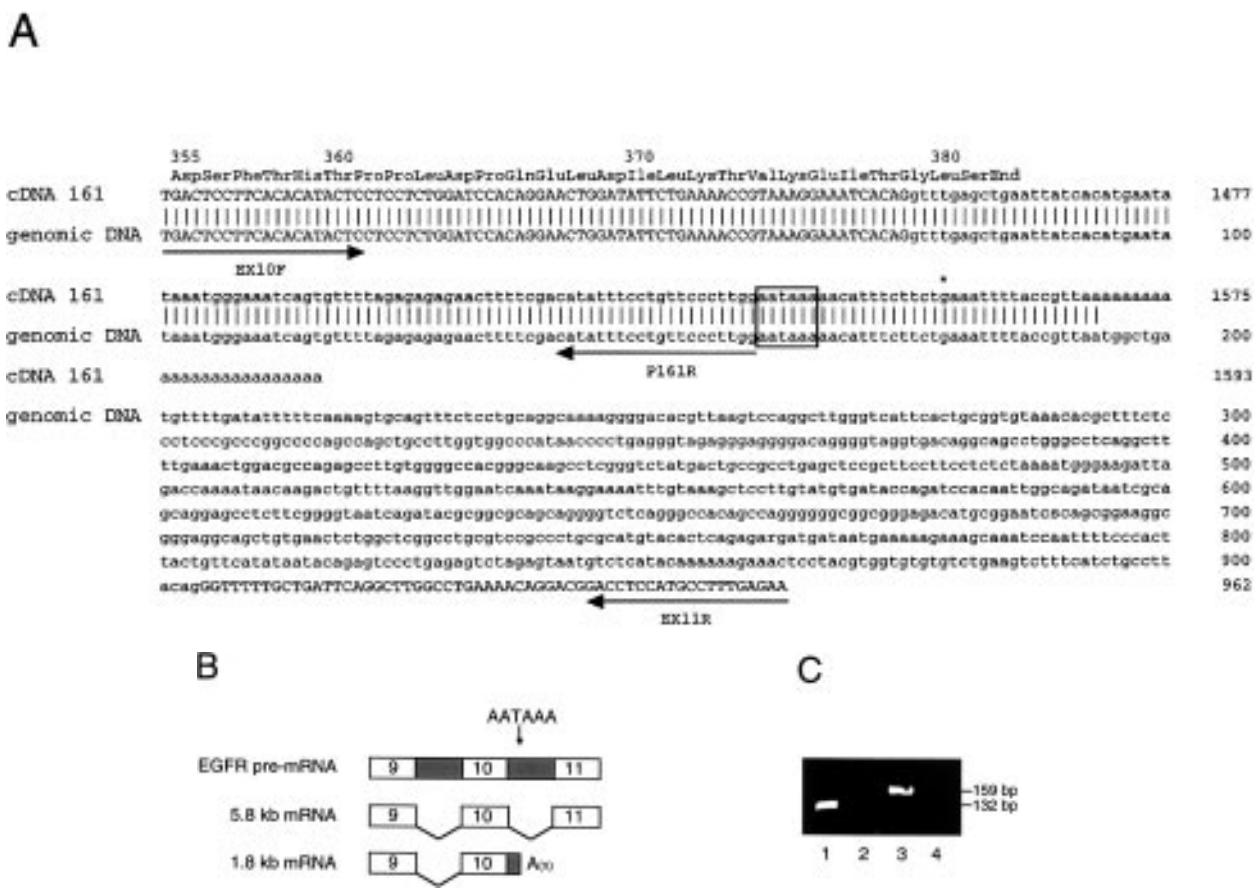

Figure 3. Generation and expression of the 1.8 kb alternative EGFR transcript. (**A**) Sequence comparison of cDNA 161 from position 1378 with genomic sequence of the EGFR locus. The DNA sequence from clone 161 is shown at the top with the deduced amino acid sequence of the encoded protein depicted above. The genomic sequence is shown below, and was derived from a PCR product amplified from human DNA with the primers EX10F and EX11R (arrows) corresponding to exons 10 and 11 respectively. The PCR primer specific for the 3′ untranslated region of cDNA 161 (P161R) is also shown. Exon sequences are indicated by capital letters. The vertical bars indicate identical bases. The AATAAA sequence is boxed. The poly(A) addition site in cDNA 763 is marked with an asterisk. (**B**) Partial map of the EGFR locus depicting the splicing events which generate the full-length and truncated mRNAs. The open boxes represent exons and the shaded boxes are introns. The use of an alternative polyadenylation signal in the pre-mRNA is responsible for the generation of the 1.8 kb EGFR transcript. (**C**) RT–PCR analysis of EGFR transcripts. RNA was reverse-transcribed with an oligo-dT primer. First strand cDNA was amplified by PCR using primers EX10F and EX11R (lanes 1 and 2) or with primers EX10F and P161R (lanes 3 and 4). Lanes 2 and 4 are negative controls where the reverse transcriptase was omitted.

ErbB1-S is secreted by transfected fibroblasts

Because clone 161 contains a signal peptide, but lacks the transmembrane domain, we predicted that the protein encoded by this cDNA should be secreted. To test this hypothesis, a quail fibroblast cell line, QT6, was transfected with the pDR2 mammalian expression vector containing clone 161 (pDR161) under the control of the Rous Sarcoma Virus LTR (19). Cells and conditioned media were subsequently analyzed for expression of this truncated protein by immunoprecipitation with monoclonal antibodies directed against the extracellular domain of the EGFR. Immunoprecipitation of mock-transfected cells failed to reveal a specific EGFR related protein in either cell lysates or in conditioned media, while a 115 kDa soluble, truncated EGFR was immunoprecipitated from the media of control A431 cells (Fig. 4A). As predicted, immunoprecipitation of conditioned media from transfected cells revealed a heterogeneous 55–65 kDa species that was specifically recognized by the EGFR monoclonal antibody, R1 (31).

In contrast, the anti-EGFR monoclonal antibody LA22 (32) recognized a ladder of proteins in QT6-transfected cells of approximately 42.5, 45, 48, 50, 53, 57, and 60 kDa, consistent with the predicted mass of ErbB1-S (Fig. 4B). Surprisingly, LA22 recognized the secreted EGFR product from A431 cells but did not recognize ErbB1-S in conditioned media of transfected cells. These results suggest there may be differences in protein folding between the soluble, truncated A431 EGFR and mature ErbB1-S.

We also tested anti-EGFR antibodies LA1, 225, and 528 (33,34) against ErbB1-S. While each of these antibodies immunoprecipitated the secreted 115 kDa EGFR from A431 cells, they did not recognize ErbB1-S from transfected cells (data not shown). These results suggest that the epitopes for these antibodies are in the distal half of the EGFR extracellular domain (carboxyl-terminal half of subdomain III or subdomain IV), consistent with the mapping of antibody binding sites in the chicken/human EGFR chimera (8).

DISCUSSION

In this study, we report the molecular cloning of a cDNA clone encoding a truncated, secreted form of the human EGFR. This cDNA clone contains a 1.6 kb insert and assuming a poly(A) tail of ~200 bp, is of a size consistent with a 1.8 kb transcript from the

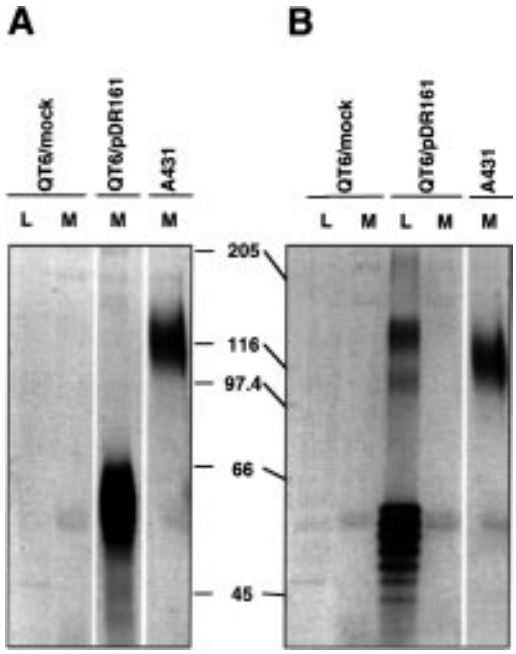

Figure 4. Metabolic labeling of soluble ErbB1 expressed in QT6 cells. QT6 cells were mock-transfected or transfected with pDR161 expressing ErbB1-S. Cells were labeled 48 h post-transfection with ^{35}S-methionine. Labeled media (M) or cell lysates (L) from QT6 cells and from positive control A431 cells were immunoprecipitated with a human EGFR extracellular domain specific monoclonal antibody, R1 (**A**), or LA22 (**B**). Immunoprecipitated samples were analyzed by SDS–PAGE on 10% gels. The positions of molecular mass standards are indicated.

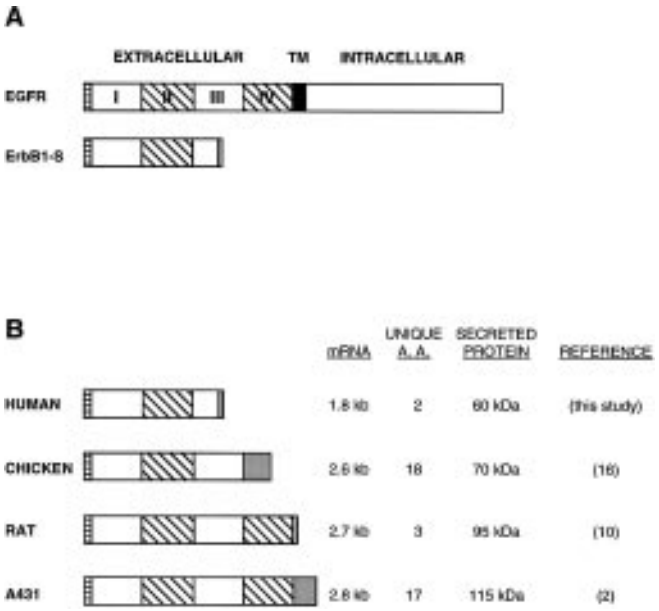

Figure 5. Schematic representation of the different EGFR forms. (**A**) A structural comparison of the human EGFR with the ErbB1-S polypeptide. (**B**) A structural comparison of the truncated receptors encoded by the human, chicken, rat, and the rearranged A431 EGFR locus is illustrated. The transcript sizes giving rise to the soluble receptors, the number of unique amino acids in the soluble receptor, and the observed size of the mature polypeptides are also listed. Horizontally hatched boxes, signal peptides; diagonally hatched boxes, cysteine-rich domains; stippled boxes, unique carboxy-terminal residues; I to IV, subdomains within the extracellular ligand-binding domain.

human EGFR gene. We further demonstrate that this transcript is expressed in human placenta and that it arises by the read-through of a splice donor site and the use of an alternative poly(A) addition signal located in intron 10. In addition, translation of this cDNA in transfected fibroblasts produces a secreted 60 kDa protein that can be immunoprecipitated with a monoclonal antibody specific for the extracellular domain of the EGFR.

It is not known what factors might be involved in the generation of variant EGFR transcripts. Conceivably, proteins may be involved in promoting or inhibiting either the splicing or the cleavage-polyadenylation reactions. Another possibility is that the variant transcripts may initiate at different transcription start sites. The human EGFR promoter region does not possess typical TATA or CAAT boxes, and RNA transcription has been shown previously to initiate at multiple sites (27,28). The 5′ sequence of cDNA 161 extends 244 nucleotides from the translation start site, suggesting that this transcript was initiated from the major *in vivo* start site located at position –255 (27). Thus, the selective use of polyadenylation signals in the expression of ErbB1-S does not appear to be associated with the differential use of transcription start sites.

Human ErbB1-S contains extracellular domains I, II, and the amino-terminal half of subdomain III in the EGFR extracellular domain (Fig. 5A) and resembles the secreted receptor encoded by the avian 2.6 kb c-*erb*B1 transcript (Fig. 5B). However, human ErbB1-S is structurally distinct from the truncated EGFR produced in normal rats and the human A431 carcinoma cell line. Both the rat and the A431 truncated receptors contain subdomain IV and diverge from the full-length receptor five amino acids upstream of the transmembrane domain.

The soluble product of the avian c-*erb*B1 gene is fully capable of binding to human transforming growth factor alpha (TGFα) *in vitro* (16). However, an important, yet unanswered question regarding the potential function of human ErbB1-S is whether it is able to bind to ligand. If ErbB1-S is capable of binding to ligand, it is possible that there will be both quantitative and qualitative differences compared with the EGFR. The 170 kDa EGFR binds to members of the EGF family, including EGF, TGFα, amphiregulin, betacellulin, and heparin-binding EGF-like growth factor. Although EGF and TGFα bind to the human EGFR with similar affinities, and are able to compete with each other for binding (7,35), they do not appear to bind identically (36,37). Therefore, ErbB1-S may be capable of binding to a subset of EGF family members, thereby modulating the activity of the transmembrane receptor.

Regardless of ligand-binding ability, ErbB1-S may be able to form inactive heterodimers with the cell surface EGFR, and thereby interfere with signal transduction. Indeed, soluble EGFR from A431 cells has been demonstrated to inhibit tyrosine kinase activity of the transmembrane receptor (38). In these *in vitro* studies, inhibition did not appear to be the result of competition for ligand-binding, but rather was proposed to occur through a direct interaction of the soluble receptor with the 170 kDa EGFR. Moreover, the avian soluble c-*erb*B1 product has been demonstrated to block TGFα-dependent soft agar colony formation in chicken embryo fibroblast cells, further demonstrating the inhibitory potential of ErbB1-S (16).

Interestingly, a soluble EGFR related inhibitor (ERI, also called astrocyte mitogen inhibitor) has been identified in normal rat brain extracts that cross-reacts with anti-EGFR antibodies (39,40). A reduction in the levels of soluble ERI following injury correlates with astrocyte mitogenic activity and a marked increase in astrocyte membrane EGFR. In another study, brain lesions significantly increased the level of the mRNA encoding the 170 kDa EGFR, but did not affect the much less abundant 2.7 kb transcript encoding the secreted EGFR (41). Together, these results imply that brain ERI represents a soluble form of the EGFR that is the product of an alternatively spliced transcript and its expression in brain is correlated with astrocyte proliferation occurring in response to injury (39,40).

In addition, soluble EGF receptors may play a role in embryonic development. Indeed, the level of the 2.7 kb EGFR transcript was significantly elevated relative to the full-length mRNA in day 10 rat decidua (10). Furthermore, a recent study focusing on the tissue- and stage-specificity of alternatively spliced EGFR transcripts in the rat incisor, demonstrated that the soluble EGFR is expressed predominantly in epithelial supra-ameloblastic cells during the secretion stage, while expression of the transmembrane receptor is restricted to the maturation stage (42).

In summary, these data demonstrate that an alternatively spliced transcript from the EGFR gene produces a secreted 60 kDa protein. Although the function of ErbB1-S is presently unclear, it may have a role in regulating cell proliferation and differentiation. Additional studies are in progress to determine the function of this naturally occurring soluble receptor in normal tissue and its potential growth inhibitory role in malignant tissues.

ACKNOWLEDGEMENTS

We gratefully acknowledge A. Lampland for assistance with the RT–PCR assay. This work was supported by NIH grants CA09441 and CA68747 to J.L.R. and CA57534 to N.J.M. and by the Mayo Foundation.

REFERENCES

1 Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. and Waterfield, M. D. (1984) *Nature*, **307,** 521–527.
2 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.
3 Coussens, L., Yang-Feng, T. L., Liao, Y.-C., Chen, E., Gray, A., McGrath, J., Seeburg, P. H., Libermann, T. A., Schlessinger, J., Francke, U., Levinson, A. and Ullrich, A. (1985) *Science*, **230,** 1132–1139.
4 Kraus, M. H., Issing, W., Miki, T., Popescu, N. C. and Aaronson, S. A. (1989) *Proc. Natl Acad. Sci. USA*, **86,** 9193–9197.
5 Plowman, G. D., Whitney, G. S., Neubauer, M. G., Green, J. M., McDonald, V. L., Todaro, G. J. and Shoyab, M. (1990) *Proc. Natl Acad. Sci. USA*, **87,** 4905–4909.
6 Plowman, G. D., Culouscou, J.-M., Whitney, G. S., Green, J. M., Carlton, G. W., Foy, L., Neubauer, M. G. and Shoyab, M. (1993) *Proc. Natl Acad. Sci. USA*, **90,** 1746–1750.
7 Lax, I., Johnson, A., Howk, R., Sap, J., Bellot, F., Winkler, M., Ullrich, A., Vennstrom, B., Schlessinger, J. and Givol, D. (1988) *Mol. Cell. Biol.*, **8,** 1970–1978.
8 Lax, I., Fischer, R., Ng, C., Segre, J., Ullrich, A., Givol, D. and Schlessinger, J. (1991) *Cell Regul.*, **2,** 337–345.
9 Maihle, N. J., Flickinger, T. W., Raines, M. A., Sanders, M. L. and Kung, H.-J. (1991) *Proc. Natl Acad. Sci. USA*, **88,** 1825–1829.
10 Petch, L. A., Harris, J., Raymond, V. W., Blasband, A., Lee, D. C. and Earp, H. S. (1990) *Mol. Cell. Biol.*, **10,** 2973–2982.
11 Merlino, G. T., Ishii, S., Whang-Peng, J., Knutsen, T., Xu, Y.-h., Clark, A. J. L., Stratton, R. H., Wilson, R. K., Ma, D. P., Roe, B. A., Hunts, J. H., Shimizu, N. and Pastan, I. (1985) *Mol. Cell. Biol.*, **5,** 1722–1734.
12 Hunts, J. H., Shimizu, N., Yamamoto, T., Toyoshima, K., Merlino, G. T., Xu, Y.-h. and Pastan, I. (1985) *Somat. Cell Mol. Genet.*, **11,** 477–484.
13 Scott, G. K., Robles, R., Park, J. W., Montgomery, P. A., Daniel, J., Holmes, W. E., Lee, J., Keller, G. A., Li, W.-L., Fendly, B. M., Wood, W. I., Shepard, H. M. and Benz, C. C. (1993) *Mol. Cell. Biol.*, **13,** 2247–2257.
14 Katoh, M., Yazaki, Y., Sugimura, T. and Terada, M. (1993) *Biochem. Biophys. Res. Commun.*, **192,** 1189–1197.
15 Heaney, M. L. and Golde, D. W. (1993) *Blood*, **82,** 1945–1948.
16 Flickinger, T. W., Maihle, N. J. and Kung, H.-J. (1992) *Mol. Cell. Biol.*, **12,** 883–893.
17 Prywes, R., Livneh, E., Ullrich, A. and Schlessinger, J. (1986) *EMBO J.*, **5,** 2179–2190.
18 Chen, W. S., Lazar, C. S., Poenie, M., Tsien, R. Y., Gill, G. N. and Rosenfeld, M. G. (1987) *Nature*, **328,** 820–823.
19 Murphy, A. J. M., Kung, A. L., Swirski, R. A. and Schimke, R. T. (1992) *Methods*, **4,** 111–131.
20 Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.*, **162,** 156–159.
21 Moscovici, C., Moscovici, M. G., Jiminez, H., Lai, M. M. C., Hayman, M. J. and Vogt, P. K. (1977) *Cell*, **11,** 95–103.
22 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1994) Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York.
23 Maihle, N. J., Raines, M. A., Flickinger, T. W. and Kung, H.-J. (1988) *Mol. Cell. Biol.*, **8,** 4868–4876.
24 Ilekis, J. V., Stark, B. C. and Scoccia, B. (1995) *Mol. Reprod. Dev.*, **41,** 149–156.
25 Chen, E. Y., Liao, Y.-C., Smith, D. H., Barrera-Saldaña, H. A., Gelinas, R. E. and Seeburg, P. H. (1989) *Genomics*, **4,** 479–497.
26 Johnson, A. C., Ishii, S., Jinno, Y., Pastan, I. and Merlino, G. T. (1988) *J. Biol. Chem.*, **263,** 5693–5699.
27 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.
28 Haley, J., Whittle, N., Bennett, P., Kinchington, D., Ullrich, A. and Waterfield, M. (1987) *Oncogene Res.*, **1,** 375–396.
29 Reiter, J. L. and Maihle, N. J. (unpublished) cDNA 161, U48722; intron 10, U48723; exon 4/intron 4, U48724; intron 5/exon 6, U48725; exon 8/intron 8/exon 9, U48726; intron 11/exon 12, U48727; intron 12/exon 13/intron 13, U48728; and exon 15/intron 15, U48729.
30 Callaghan, T., Antczak, M., Flickinger, T., Raines, M., Myers, M. and Kung, H.-J. (1993) *Oncogene*, **8,** 2939–2948.
31 Waterfield, M. D., Mayes, E. L. V., Stroobant, P., Bennet, P. L. P., Young, S., Goodfellow, P. N., Banting, G. S. and Ozanne, B. (1982) *J. Cell. Biochem.*, **20,** 149–161.
32 Wu, D., Wang, L., Sato, G. H., West, K. A., Harris, W. R., Crabb, J. W. and Sato, J. D. (1989) *J. Biol. Chem.*, **264,** 17469–17475.
33 Kawamoto, T., Sato, J. D., Le, A., Polikoff, J., Sato, G. H. and Mendelsohn, J. (1983) *Proc. Natl Acad. Sci. USA*, **80,** 1337–1341.
34 Sato, J. D., Kawamoto, T., Le, A. D., Mendelsohn, J., Polikoff, J. and Sato, G. H. (1983) *Mol. Biol. Med.*, **1,** 511–529.
35 Massagué, J. (1983) *J. Biol. Chem.*, **258,** 13614–13620.
36 Winkler, M. E., O'Connor, L., Winget, M. and Fendly, B. (1989) *Biochemistry*, **28,** 6373–6378.
37 Harte, M. T. and Gentry, L. E. (1995) *Arch. Biochem. Biophys.*, **322,** 378–389.
38 Basu, A., Raghunath, M., Bishayee, S. and Das, M. (1989) *Mol. Cell. Biol.*, **9,** 671–677.
39 Nieto-Sampedro, M. (1988) *Science*, **240,** 1784–1786.
40 Nieto-Sampedro, M. and Broderick, J. T. (1989) *J. Neurosci. Res.*, **22,** 28–35.
41 Junier, M.-P., Hill, D. F., Costa, M. E., Felder, S. and Ojeda, S. R. (1993) *J. Neurosci.*, **13,** 703–713.
42 Davideau, J.-L., Sahlberg, C., Blin, C., Papagerakis, P., Thesleff, I. and Berdal, A. (1995) *Int. J. Dev. Biol.*, **39,** 605–615.
43 Kagayama, R., Merlino, G. T. and Pastan, I. (1988) *J. Biol. Chem.*, **263,** 6329–6336.