Reports

Human Endothelial Cell Growth Factor: Cloning, Nucleotide Sequence, and Chromosome Localization

MICHAEL JAYE,* RICHARD HOWK, WILSON BURGESS, GEORGE A. RICCA, ING-MING CHIU, MARK W. RAVERA, STEPHEN J. O'BRIEN, WILLIAM S. MODI, THOMAS MACIAG, WILLIAM N. DROHAN

Several of the endothelial cell polypeptide mitogens that have been described probably play a role in blood vessel homeostasis. Two overlapping complementary DNA clones encoding human endothelial cell growth factor (ECGF) were isolated from a human brain stem complementary DNA library. Southern blot analysis suggested that there is a single copy of the ECGF gene and that it maps to human chromosome 5 at bands 5q31.3 to 33.2. A 4.8-kilobase messenger RNA was present in human brain stem messenger RNA. The complete amino acid sequence of human ECGF was deduced from the nucleic acid sequence of these clones; it encompasses all the well-characterized acidic endothelial cell polypeptide mitogens described by several laboratories. The ECGF-encoding open reading frame is flanked by translation stop codons and provides no signal peptide or internal hydrophobic domain for the secretion of ECGF. This property is shared by human interleukin-1, which is approximately 30 percent homologous to ECGF.

ARTERIES, VEINS, AND CAPILLARIES in vivo are protected from thrombogenesis by the presence of the endothelial cell (*1*). The endothelium exists as a single cell layer, and the cells generally do not show mitosis except after wounding. Factors that mediate endothelial cell function are important to the maintenance and regulation of blood vessel homeostasis (*2*) and may be important in the pathobiology of tumor growth and development, the genesis of atherosclerosis, and events related to the aging process.

Specific modifiers of endothelial cell migration and proliferation include the extracellular matrix, polypeptide mitogens, hydrolases, and heparin and have a role in the process of neovascularization in vivo. The diversity of these and other effectors has yielded the cascade hypothesis, which assumes the existence of pathways that can act either independently or synergistically to generate the neovascularization process (*3*).

Two forms of endothelial cell growth factor (ECGF) have been isolated from bovine neural tissue and termed α- and β-ECGF (*4*). These polypeptides are members of a family of anionic endothelial cell polypeptide mitogens (*5*) that also includes acidic fibroblast growth factor 1 (aFGF-1) (*6*) and acidic fibroblast growth factor 2 (aFGF-2) (*7*), eye-derived growth factor II (*8*), and α-retina–derived growth factor (*9*). The biological activity of ECGF is mediated by a high-affinity polypeptide receptor present on the endothelial cell surface (*10, 11*) and is potentiated by a structural interaction with the glycosaminoglycan heparin (*4, 10, 12*). Polypeptides of the ECGF family are potent modulators of the angiogenic response in vivo (*7, 13, 14*). The cationic or basic fibroblast growth factor family of endothelial cell polypeptide mitogens has also been described (*9, 15–17*). These polypeptides share many biochemical features with the polypeptides of the ECGF family, such as similar molecular weights (approximately 17,000), affinity for heparin, and angiogenic activity in vivo. Although it is unclear whether the ECGF and basic FGF families of polypeptide mitogens mediate separate angiogenic pathways or act synergistically, the isolation of two angiogenic polypeptide families from diverse sources suggests that the mechanism of neovascularization in vivo is complex. To further characterize the ECGF family of polypeptide mitogens and to facilitate studies correlating polypeptide structure with biological function, we isolated complementary DNA (cDNA) clones that encode the polypeptide human ECGF and determined the primary structure of the endothelial cell polypeptide mitogen by nucleotide sequence analysis.

A cDNA library was constructed in λgt11 (*18*) by using polyadenylated RNA extracted from human brain stem (*19*). The cDNA library, which consists of 5.7×10^6 phage, of which approximately 65 percent are recombinant, was screened for clones containing ECGF inserts with a synthetic oligonucleotide probe. The oligonucleotide was designed with amino acid sequence data obtained from automated Edman degradation of a trypsin-derived fragment that is common to both α- and β-ECGF (*20*). The amino acid sequence of this peptide is Ile-Leu-Pro-Asp-Gly-Thr-Val-Asp-Gly-Thr-Lys. Instead of using a mixture of oligonucleotides containing all of the possible coding sequences (owing to the degeneracy of the genetic code), we designed a long unique oligonucleotide. Such oligonucleotide probes have been shown to be successful in screening complex cDNA (*21*) and genomic (*22*) libraries.

The following criteria were used in the construction of the oligonucleotide probe: (i) the dinucleotide CG was avoided because it is underrepresented in eukaryotic DNA (*23*), (ii) preferred codons (*24*) were used wherever possible, and (iii) unusual base pairing was allowed because of the natural occurrence of G·T, I·T, I·A, and I·C base pairs, which occur in the interaction between transfer RNA anticodons and messenger RNA (mRNA) codons (*25*). The use of inosine (I) in the construction of a hybridization probe was first demonstrated by Ohtsuka *et al.* (*26*). The 33-base oligonucleotide used to screen the human brain stem cDNA library contained six inosines and had the nucleotide sequence 5′-ATTTTTC-CIGATGGIACIGTIGATGGIACIAAA-3′. Additional amino acid sequences were derived from automated Edman degradation of undigested α-ECGF and cyanogen bromide–digested α- and β-ECGF (*20*). These sequences were used to design two 48-base oligonucleotides according to the strategy described above.

The human brain stem cDNA library (1.5×10^6 plaques) was screened with the ^{32}P-labeled 33-base oligonucleotide. Two plaques generated a positive autoradiographic signal. The clones were purified to homogeneity by repeated hybridization with the 33-base oligonucleotide. These clones, designated clone 1 and clone 29, were analyzed by Eco RI digestion and revealed cDNA inserts of 2.2 and 0.2 kilobases (kb), respectively. Nick translation of the cloned cDNA and its subsequent use as a radiolabeled probe in Southern blot analysis (*27*) demonstrated that clones 1 and 29

M. Jaye, G. A. Ricca, I.-M. Chiu, M. W. Ravera, W. N. Drohan, Molecular Biology Division, Meloy Laboratories, Springfield, VA 22151.
R. Howk, W. Burgess, T. Maciag, Department of Cell Biology, Biotechnology Research Center, Rockville, MD 20850.
S. J. O'Brien and W. S. Modi, Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick Cancer Research Center, Frederick, MD 21701.

*To whom correspondence should be addressed.

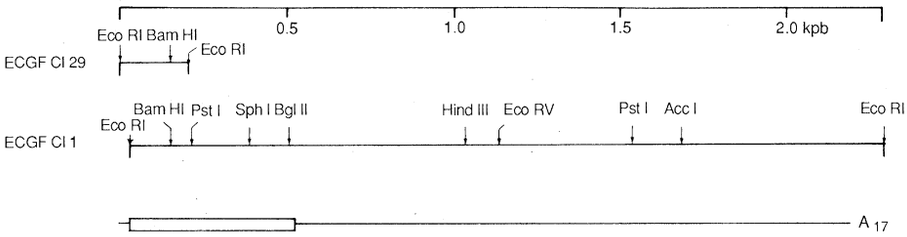

Fig. 1. Schematic diagram of human ECGF cDNA clones. Construction of the human brain stem cDNA library in λgt11 has been described (*19*). Phage (1.5×10^6) of the human brain stem cDNA library were plated on a lawn of *Escherichia coli* Y1088 [supE supF metB trpR hsdR$^-$ hsdM$^+$ tonA21 strA lacU169 (proC::Tn5) (pMC9)] and incubated at 42°C for 6 hours. Phage DNA was transferred to duplicate nitrocellulose filters and denatured (*27*), and the filters were prehybridized at 42°C in 6× SSPE [1× SSPE is 0.18*M* NaCl, 0.01*M* $NaPO_4$ (*p*H 7.2), and 0.001*M* EDTA], 2× Denhardt's (1× Denhardt's is 0.2% each of Ficoll, polyvinylpyrrolidone, and bovine serum albumin), 5% dextran sulfate, and denatured salmon sperm DNA (100 μg/ml). Approximately 30 pmol of the 33-base oligonucleotide was labeled by incubation with [^{32}P]adenosine triphosphate and T_4 polynucleotide kinase (*27*). Following 4 hours of prehybridization, the ^{32}P-labeled oligonucleotide was added, and hybridization continued overnight at 42°C. Unhybridized probe was removed by sequential washing at 37°C in 2× SSPE and 0.1% sodium dodecyl sulfate (SDS). The hybridized probe was detected after overnight autoradiography. The two clones that were isolated, ECGF clones 1 and 29, were purified to homogeneity by repeated cycles of purification with the 33-base oligonucleotide as a hybridization probe. The open box represents the open reading frame encoding human ECGF. The Eco RI sites correspond to synthetic oligonucleotide linkers used in the construction of the cDNA library. The polyadenylate tail at the 3′ end of clone 1 is shown as A_{17}.

were overlapping clones. Hybridization with both radiolabeled 48-base oligonucleotides was observed with clone 1, whereas clone 29 failed to hybridize with the 48-base oligonucleotide derived from the amino acid sequence of cyanogen bromide–digested α- and β-ECGF. A restriction enzyme map and a schematic diagram of the cDNA inserts of clones 1 and 29 is shown in Fig. 1.

The cDNA inserts from clones 1 and 29 were isolated, subcloned into M13mp18, and the ECGF-encoding open reading frame and flanking regions were sequenced by the chain termination method (*28*). Examination of the nucleotide sequence revealed an open reading frame of 465 nucleotides encoding human ECGF (Fig. 2). The amino acid sequence Ile-Leu-Pro-Asp-Gly-Thr-Val-Asp-Gly-Thr-Lys used to design the oligonucleotide probe was encoded by nucleotides 118 through 150 (Fig. 2). The 155 amino acids encoding human ECGF were flanked by translation stop codons.

The complete amino acid sequence of bovine aFGF-1 has been reported (*29*) and agrees well with the amino acid sequence deduced from the human ECGF cDNA sequence. However, bovine aFGF-1 begins with the amino-terminal sequence Phe-Asn-Leu-Pro-Leu and thus is 15 amino acids smaller than human ECGF (Fig. 3, line II). The predicted amino-terminal modification of human ECGF that would result in the formation of aFGF-1 involves the processing of a Lys-Phe bond (Fig. 3, residues 15 and 16), a cleavage that may involve serine proteases.

Amino acid sequence determination at the amino terminus of bovine α-ECGF (*20*) or aFGF-2 (*7*) reveals the sequence Asn-Tyr-Lys-Lys-Pro . . . , corresponding to a polypeptide that is 21 amino acids shorter at the amino terminus than human ECGF. Digestion of bovine α- and β-ECGF with trypsin reveals two tryptic peptides unique to β-ECGF (*4*). These two tryptic peptides may be predicted from the amino acid sequence of human ECGF (amino acids 1 to 15 and 16 to 24) (Fig. 2). Thus, our human ECGF sequence most likely represents that of human β-ECGF. However, until a precursor-product relation between β-ECGF and the smaller polypeptides aFGF-1, aFGF-2, and α-ECGF is established experimentally, these alignments do not eliminate the alternative possibilities of separate genes encoding the acidic endothelial cell mitogens or alternative splicing of a precursor mRNA moiety.

The formation of α-ECGF or aFGF-2 from β-ECGF is noteworthy because it involves the processing of a Gly-Asn bond (Fig. 3, residues 21 and 22), which should be refractory to trypsin-like enzymes. Analysis of the amino-terminal sequence of bovine basic FGF isolated from various tissues (*16*) suggests that a Gly-His bond that aligns with the Gly-Asn bond in β-ECGF (Fig. 3) is similarly cleaved. Neural tissue may contain a unique proteolytic activity with the ability to cleave these polypeptides at these specific peptide bonds.

Comparison of the complete amino acid

```
                                                                    -1
                                                GAATTCGGGAACGCGCCACAAGCAGCAGCTGCTGAGCC
                                                                    *

1                                                                  120
ATGGCTGAAGGGGAAATCACCACCTTCACAGCCCTGACCGAGAAGTTTAATCTGCCTCCAGGGAATTACAAGAAGCCCAAACTCCTCTACTGTAGCAACGGGGGCCACTTCCTGAGGATC
M  A  E  G  E  I  T  T  F  T  A  L  T  E  K  F  N  L  P  P  G  N  Y  K  K  P  K  L  L  Y  C  S  N  G  G  H  F  L  R  I

                                                                   240
CTTCCGGATGGCACAGTGGATGGGACAAGGGACAGGAGCGACCAGCACATTCAGCTGCAGCTCAGTGCGGAAAGCGTGGGGGAGGTGTATATAAAGAGTACCGAGACTGGCCAGTACTTG
L  P  D  G  T  V  D  G  T  R  D  R  S  D  Q  H  I  Q  L  Q  L  S  A  E  S  V  G  E  V  Y  I  K  S  T  E  T  G  Q  Y  L

                                                                   360
GCCATGGACACCGACGGGCTTTTATACGGCTCACAGACACCAAATGAGGAATGTTTGTTCCTGGAAAGGCTGGAGGAGAACCATTACAACACCTATATATCCAAGAAGCATGCAGAGAAG
A  M  D  T  D  G  L  L  Y  G  S  Q  T  P  N  E  E  C  L  F  L  E  R  L  E  E  N  H  Y  N  T  Y  I  S  K  K  H  A  E  K

                                                                   480
AATTGGTTTGTTGGCCTCAAGAAGAATGGGAGCTGCAAACGCGGTCCTCGGACTCACTATGGCCAGAAAGCAATCTTGTTTCTCCCCCTGCCAGTCTCTTCTGATTAAAGAGATCTGTTC
N  W  F  V  G  L  K  K  N  G  S  C  K  R  G  P  R  T  H  Y  G  Q  K  A  I  L  F  L  P  L  P  V  S  S  D  *

                                                                   600
TGGTGTTGACCACTCCAGAGAAGTTTCGAGGGGTCCTCACCTGGTTGACCCCAAAAATGTTCCCTTGACCATTGGCTGCGCTAACCCCCAGCCCACAGAGCCTGAATTTGTAAGCAACTT
```

Fig. 2. Nucleotide and deduced amino acid sequence of human β-ECGF. The cDNA inserts from ECGF clones 1 and 29 were subcloned into M13mp18, and the ECGF-encoding open reading frame and flanking regions were sequenced by the chain termination method (*28*). In-frame stop codons at the 5′ and 3′ ends of the ECGF-encoding open reading frame are indicated by the asterisks. The single-letter notation for amino acids is used: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

```
        1        10        20        30        40        50        60
I.      MAEGEITTFT ALTEKFNLPP GNYKKPKLLY CSNGGHFLRI LPDGTVDGTR DRSDQHIQLQ
II.                ****L     ********** *****Y**** *********K **********
III.         PALP EDGGSGAF** *HF*D**R** *K***F**** H***R***V* EK**P**K**

        70        80        90        100       110       120
I.      LSAESVGEVY IKSTETGQYL AMDTDGLLYG SQTPNEECLF LERLEENHYN TYISKKHAEK
II.     *C***I**** ********F* ********** ********** ********** **********
III.    *Q**ER*V*S **GVCANR** **KE**R*LA *KCVTD**F* F****S*N** **R*R*--YS

        130       140       150   155
I.      NWFVGLKKNG SCKRGPRTHY GQKAILFLPL PVSSD
II.     H********* RS*L*****F ********** *****
III.    S*Y*A**RT* QY*L**K*GP *********M SAK*
```

Fig. 3. Homology of human ECGF with bovine acidic fibroblast growth factor, and bovine basic fibroblast growth factor. (Line I) The deduced human ECGF amino acid sequence derived from ECGF clones 1 and 29; (line II) amino acid sequence of bovine acidic fibroblast growth factor (*29*); (line III) amino acid sequence of bovine basic fibroblast growth factor (*16*). The following symbols are used: *, identical amino acid; –, a gap has been introduced in the amino acid sequence to allow maximum alignment.

sequence of bovine basic FGF (*16*) with the amino acid sequence of human β-ECGF indicates an overall sequence identity of 50.7 percent (Fig. 3, line III). Clusters of basic amino acids are conserved among bovine basic FGF and human ECGF; these may represent the site of heparin binding (*30*) to these growth factors. However, the most significant homology between bovine basic FGF and human ECGF occurs near the respective carboxyl termini, where a stretch of nine amino acids with the sequence Gly-Gln-Lys-Ala-Ile-Leu-Phe-Leu-Pro is perfectly conserved in both polypeptides. The significance of this homology in terms of structure and function is not known. In contrast, the amino-terminal region (residues 7 to 18) of human ECGF contains no homology with basic FGF. The short, imperfect repeats of the amino acid sequence observed in bovine basic FGF also show no conservation in human ECGF.

Tetrapeptides that are related to the sequence NH_2-Arg-Gly-Asp-Ser-COOH are thought to be involved in the cellular attachment of fibronectin, fibrinogen, and von Willebrand factor (*31*). As originally noted by Esch *et al.* (*16*), bovine basic FGF contains multiple sequences of related structure that may mediate interactions with cellular surfaces. These or related tetrapeptides do not occur in the deduced amino acid sequence of human ECGF.

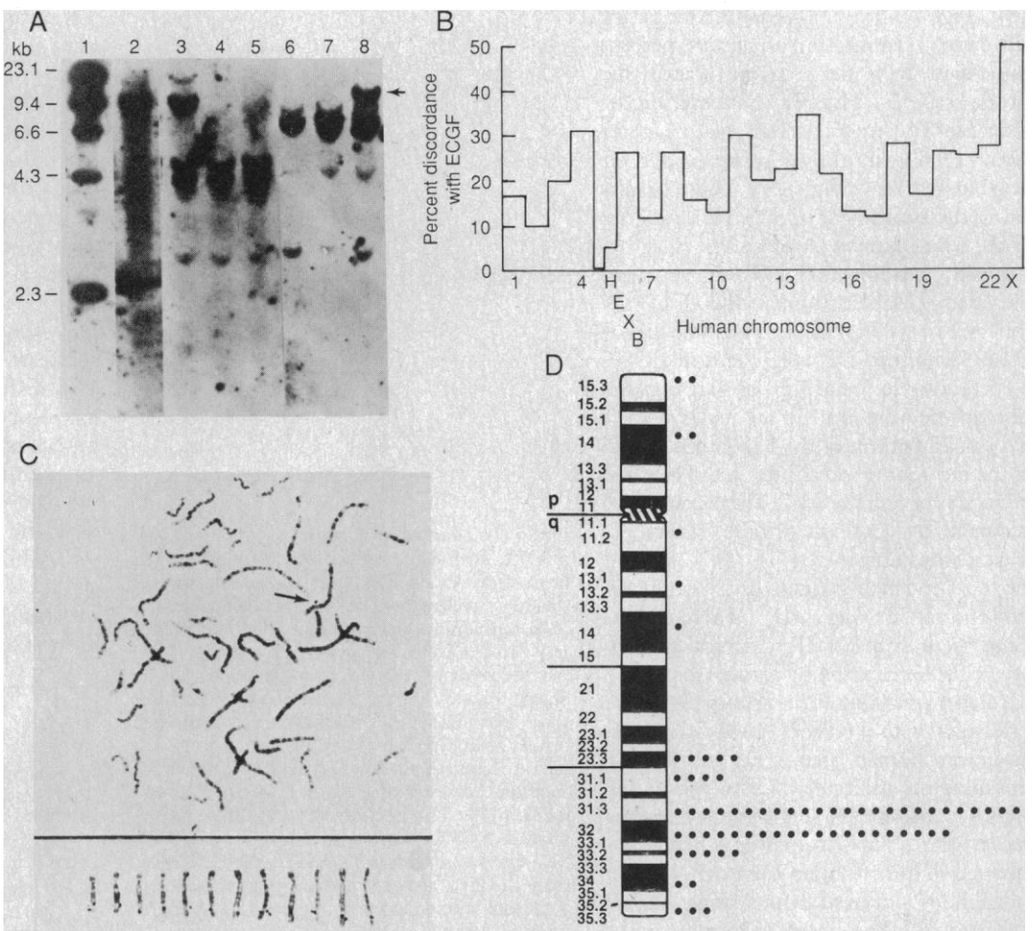

Fig. 4. (A) Hybridization of the labeled ECGF cDNA clone 1 probe to Eco RI digests of human, RAG mouse, E36 Chinese hamster, and selected hybrids. Restriction endonuclease–digested genomic DNA was separated in agarose gels, transferred to GeneScreen Plus (New England Nuclear), and hybridized to ^{32}P-labeled human ECGF clone 1 cDNA in a solution containing 0.60*M* NaCl, 0.06*M* sodium citrate, 0.2% Ficoll, 0.2% bovine serum albumin, 0.2% polyvinylpyrrolidone, 0.1% SDS, and low molecular weight salmon testes DNA (50 μg/ml) at 65°C for 18 to 29 hours. Sizes are in kilobases. (Lane 1) Hind III–digested lambda DNA molecular weight standard. (Lane 2) Human lymphocyte DNA. (Lane 3) Chinese hamster hybrid 80H7C. (Lane 4) Chinese hamster hybrid 80H8D. (Lane 5) E36 Chinese hamster. (Lane 6) RAG mouse. (Lane 7) Mouse hybrid 70M5C. (Lane 8) Mouse hybrid 70M6C. The 9.4-kb human-specific fragment is indicated by the arrow. (B) Plot of the frequency of discordancy between human ECGF sequences and each of the 23 human chromosomes in a panel of 41 somatic cell hybrids. The chromosome scores of hybrids represent the consensus of G-trypsin karyotype and isozymal scores (*32*). (C) In situ hybridization of ^{3}H-labeled ECGF probe to metaphase chromosome preparations and to several labeled chromosomes 5. (D) Idiogram of grain distribution on chromosome 5.

We sought to determine whether the acidic endothelial cell polypeptide mitogens are encoded by one or multiple homologous genes by Southern blotting. Upon digestion of either human lymphocyte or HeLa cell DNA with Eco RI, Pst I, Pvu II, or Bam HI, and with the 2.2-kb cDNA insert of human ECGF clone 1 as probe, one to three bands were revealed with an average total size of 12 ± 3 kb. The simple patterns observed are consistent with the proposition that the human genome contains only a single copy of a gene homologous to the ECGF cDNA probe. The results suggest that β-ECGF, α-ECGF, aFGF-1, and aFGF-2 are encoded by a common gene. In addition, the gene for ECGF is interrupted by introns.

The human chromosome containing the ECGF sequences was identified by a Southern transfer analysis of cellular DNA's from rodent-human somatic cell hybrids (*32*). These hybrids retain the entire rodent genome but lose human chromosomes in different combinations. Cellular DNA (approximately 10 μg) was digested with Eco RI, transferred to GeneScreen Plus membranes, and hybridized with a nick-translated probe derived from the 2.2-kb cDNA insert of human ECGF clone 1. The presence of the diagnostic human 9.4-kb fragment (Fig. 4A) was 100 percent concordant with human chromosome 5 and 98 percent concordant with the isozyme marker for chromosome 5 (HEXB), but was highly discordant (12 to 77 percent) with each of the other human chromosomes and their included markers (Fig. 4B). These results permit the assignment of the structural gene for ECGF to human chromosome 5.

In situ hybridization (*33*) of the nick-translated ^{3}H-labeled human ECGF cDNA clone to normal human chromosome preparations confirmed the assignment of ECGF to chromosome 5 and further localized the gene on the long arm. In an analysis of 81 mitoses, 25 percent of the 257 grains scored on the metaphase spreads exhibited label on the region of 5q31.3-33.2. These results are illustrated by a cell metaphase, several labeled chromosomes 5 (Fig. 4C), and an idiogram indicating grain distribution on chromosome 5 (Fig. 4D). These results assign the human ECGF structural gene to the q31.3-33.2 region of chromosome 5.

Northern blotting experiments (*34*) were performed with polyadenylated RNA derived from human brain stem and liver to determine the molecular size of the ECGF mRNA transcript. A single molecular species at 28*S*, or approximately 4.8 kb, was observed in human brain stem mRNA, but not in mRNA derived from a single human liver (Fig. 5). The observation that only a singular mRNA species is homologous with the ECGF probe suggests that the multiple acidic endothelial cell polypeptide mitogens discussed above are not derived by alternative RNA splicing. Since the β-ECGF coding sequence is flanked by translation stop codons (Fig. 2), it is likely that the 4.8-kb mRNA contains unusually long regions of untranslated sequence. Thus, unlike many growth factors, such as epidermal growth factor (*35*) and nerve growth factor (*36*), human β-ECGF is not derived by proteolytic processing of a polypeptide precursor.

The amino-terminal amino acid of human β-ECGF deduced from the cDNA sequence is methionine, which most likely serves as the translation initiation residue. The nucleotide sequence around and including this proposed translation initiation codon is 5′-GCCAUGG-3′. This sequence conforms well to the nucleotide sequence that is optimal for initiation of protein synthesis in eukaryotes, 5′-ACCAUGG-3′ (*37*). The first 15 to 20 amino-terminal residues of human β-ECGF are relatively nonhydrophobic, suggesting that it is synthesized without an amino-terminal signal peptide. Analysis of the amino acid composition of an amino-terminal peptide derived by trypsin treatment of bovine β-ECGF (*20*) shows strong homology with the first 15 amino acids deduced from the cDNA sequence of human β-ECGF. From this we conclude that the deduced amino-terminal amino acid sequence of human β-ECGF corresponds to that of the native polypeptide and that no signal sequence or other polypeptide precursor exists.

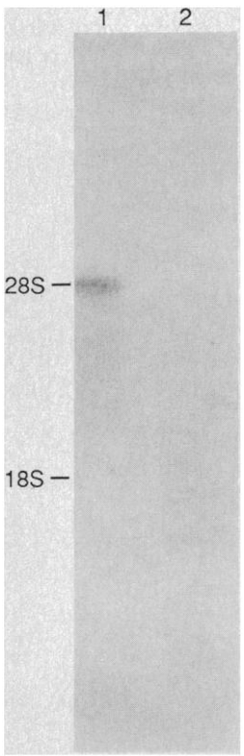

Fig. 5. Northern blot analysis of human ECGF mRNA. Ten micrograms of polyadenylated RNA from human brain stem (lane 1) and human liver (lane 2) was denatured in 2.2*M* formaldehyde and 50% formamide and fractionated by electrophoresis in a 1.25% agarose gel containing 2.2*M* formaldehyde. This was transferred to GeneScreen Plus by blotting with 10× SSPE. Blots were hybridized to ^{32}P-labeled nick-translated probes of ECGF clone 1 at 65°C for 16 hours in a mixture containing 2× SSPE, 20× Denhardt's solution, yeast transfer RNA (200 μg/ml), and 0.2% SDS. The membrane was subsequently washed at 65°C, twice with 2× SSPE and 0.2% SDS and twice with 0.2× SSPE and 0.2% SDS, then air-dried and exposed overnight to Kodak XAR film with an intensifying screen. The migration of 28*S* and 18*S* RNA is noted.

The lack of a signal peptide in human β-ECGF, a polypeptide for which high-affinity cell-surface receptors have been characterized (*10, 11*), evokes the obvious question of how it is released from the site of its biosynthesis. To our knowledge, the lack of a signal peptide or internal hydrophobic domain in human β-ECGF is shared by only one other extracellular protein—human interleukin-1 (IL-1) (*38*). Human IL-1 has approximately 30 percent homology with aFGF-1 (*7*), which, as discussed above, is most likely cleaved from β-ECGF. Auron *et al.* (*38*) speculate that IL-1β is released from stimulated monocytes or macrophages by leakage, perhaps due to cell damage. If release of ECGF is also due to leakage from cells, it is likely that neural tissue is not the sole site of synthesis of ECGF.

Baird *et al.* (*39*) demonstrated the presence of basic FGF in extracts of stimulated peritoneal macrophages. In view of the similarities between ECGF and IL-1 discussed above, as well as the similar biological activities of the basic and acidic endothelial cell polypeptide mitogens, it is reasonable to speculate that the macrophage also synthesizes ECGF. Delivery, by the stimulated macrophage, of endothelial cell polypeptide mitogens to sites of injury and inflammation may be the underlying mechanism by which these polypeptides play a role in the regulation and maintenance of homeostasis in vivo. This may also be relevant to the high-affinity binding of monocytes to endothelial cells (*40*), which appears to be an early event in the genesis of atherosclerosis (*41*).

REFERENCES AND NOTES

1. M. A. Gimbrone, Jr., *Prog. Thromb. Hemost.* **3**, 1 (1976).
2. J. Folkman, *Ann. Intern. Med.* **82**, 96 (1975).
3. T. Maciag, *Prog. Thromb. Hemost.* **7**, 167 (1984).
4. W. H. Burgess, T. Mehlman, R. Friesel, W. V. Johnson, T. Maciag, *J. Biol. Chem.* **260**, 11389 (1985).
5. A. B. Schreiber *et al.*, *J. Cell Biol.* **101**, 1623 (1985).
6. P. Böhlen, F. Esch, A. Baird, D. Gospodarowicz, *EMBO J.* **8**, 1951 (1985); F. Esch *et al.*, *Biochem. Biophys. Res. Commun.* **133**, 554 (1985).

7. K. A. Thomas *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6409 (1985).
8. J. Courty *et al.*, *Biochimie* **67**, 265 (1985).
9. P. A. D'Amore and M. Klagsbrun, *J. Cell Biol.* **99**, 1545 (1984); A. Baird, F. Esch, D. Gospodarowicz, R. Guillemin, *Biochemistry* **24**, 7855 (1985).
10. T. Maciag, T. Mehlman, R. Friesel, A. B. Schreiber, *Science* **225**, 932 (1984); A. B. Schreiber *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6138 (1985).
11. R. Friesel, W. H. Burgess, T. Mehlman, T. Maciag, *J. Biol. Chem.* **261**, 7581 (1986).
12. W. Johnson *et al.*, *J. Cell Biol.* **101**, 464 (1985); W. Johnson *et al.*, in preparation.
13. R. R. Lobb, E. M. Alderman, J. W. Fett, *Biochemistry* **24**, 4969 (1985).
14. A. B. Schreiber and T. Maciag, unpublished data.
15. D. Gospodarowicz, J. Cheng, G.-M. Lui, A. Baird, P. Böhlen, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6963 (1984).
16. F. Esch *et al.*, *ibid.*, **82**, 6507 (1985).
17. R. R. Lobb *et al.*, *J. Biol. Chem.* **261**, 1924 (1986).
18. R. A. Young and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1194 (1983); *Science* **222**, 778 (1983).
19. A human brainstem cDNA library was constructed in λgt11 by J. Kamholz, C. Puckett, and G. Ricca essentially as described [F. deFerra *et al.*, *Cell* **43**, 721 (1985)].
20. W. H. Burgess, T. Mehlman, D. R. Marshak, B. A. Fraser, T. Maciag, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
21. M. Jaye *et al.*, *Nucleic Acids Res.* **11**, 2325 (1983).
22. J. Gitschier *et al.*, *Nature (London)* **312**, 326 (1984).
23. J. Josse, A. A. Kaiser, A. Kornberg, *J. Biol. Chem.* **236**, 864 (1961).
24. R. Lathe, *J. Mol. Biol.* **183**, 1 (1985).
25. F. H. C. Crick, *ibid.* **19**, 548 (1966).
26. E. Ohtsuka, S. Matsuki, M. Ikehara, Y. Takahashi, K. Matsubara, *J. Biol. Chem.* **260**, 2605 (1985).
27. T. Maniatis, E. F. Fritch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1982).
28. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
29. G. Gimenez-Gallego *et al.*, *Science* **230**, 1385 (1985).
30. J. E. Schwartzbauer, J. W. Tamkun, I. R. Lemischka, R. O. Hynes, *Cell* **35**, 421 (1983).
31. M. D. Pierschbacher and E. Ruoslahti, *Nature (London)* **309**, 30 (1984); J. E. Sadler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6394 (1985); M. D. Pierschbacher and E. Ruoslahti, *ibid.* **81**, 5985 (1984).
32. S. J. O'Brien, T. I. Bonner, M. Cohen, C. O'Connell, W. G. Nash, *Nature (London)* **303**, 74 (1983); S. J. O'Brien, W. G. Nash, J. L. Goodwin, D. R. Lowy, E. H. Chang, *ibid.* **302**, 839 (1983); S. J. O'Brien, J. M. Simonson, M. Eichelberger, in *Techniques in Somatic Cell Genetics*, J. W. Shay, Ed. (Plenum, New York, 1982), p. 513.
33. N. R. Campbell *et al.*, *Nature (London)* **251**, 158 (1974); B. U. Zabel *et al.*, *Cytogenet. Cell Genet.* **80**, 6932 (1983).
34. M. Jaye *et al.*, *Science* **228**, 882 (1985).
35. J. Scott *et al.*, *ibid.* **221**, 236 (1983).
36. J. M. Scott *et al.*, *Nature (London)* **302**, 538 (1983).
37. M. Kozak, *Cell* **44**, 283 (1986).
38. P. E. Auron *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7907 (1984); Y. Furutani *et al.*, *Nucleic Acids Res.* **13**, 5869 (1985).
39. A. Baird, P. Mormede, P. Böhlen, *Biochem. Biophys. Res. Commun.* **126**, 358 (1985).
40. N. A. Pawlowski *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8208 (1985).
41. A. Faggioto, R. Ross, L. Harker, *Arteriosclerosis* **4**, 323 (1984); H. C. McGill, Jr., *ibid.*, p. 443.
42. We thank W. Terry for his continued support, C. Puckett for brain mRNA, and N. Gravely for expert and congenial secretarial assistance. Supported in part by grants HL23348 and AG04807 to T.M. and HL35627 to T.M. and W.N.D.

5 March 1986; accepted 13 May 1986

Nucleotide Sequence of a Bovine Clone Encoding the Angiogenic Protein, Basic Fibroblast Growth Factor

JUDITH A. ABRAHAM, AYALEW MERGIA, JACQUELINE L. WHANG, ANNETTE TUMOLO, JEFF FRIEDMAN, KATHRYN A. HJERRILD, DENIS GOSPODAROWICZ, JOHN C. FIDDES

Basic and acidic fibroblast growth factors (FGF's) are potent mitogens for capillary endothelial cells in vitro, stimulate angiogenesis in vivo, and may participate in tissue repair. An oligonucleotide probe for bovine basic FGF was designed from the nucleotide sequence of the amino-terminal exon of bovine acidic FGF, taking into account the 55 percent amino acid sequence homology between the two factors. With this oligonucleotide probe, a full length complementary DNA for basic FGF was isolated from bovine pituitary. Basic FGF in bovine hypothalamus was shown to be encoded by a single 5.0-kilobase messenger RNA; in a human hepatoma cell line, both 4.6- and 2.2-kilobase basic FGF messenger RNA's were present. Both growth factors seem to be synthesized with short amino-terminal extensions that are not found on the isolated forms for which the amino acid sequences have been determined. Neither basic nor acidic FGF has a classic signal peptide.

BASIC FIBROBLAST GROWTH FACTOR (FGF) stimulates the proliferation in vitro of a wide range of mesoderm-derived cells and, in particular, is a potent mitogen for vascular endothelial cells (*1*). It has been purified from various tissues, including pituitary, brain, hypothalamus, retina, adrenal gland, corpus luteum, and kidney (*2*). A second endothelial cell mitogen, acidic FGF, is structurally and functionally related to basic FGF but has only been detected in neural tissues such as brain, hypothalamus, and retina (*2*). Both of these growth factors have a high affinity for the glycosaminoglycan heparin (*3*) and are probably identical to several other heparin-binding endothelial cell growth factors (*4*).

In addition to being mitogenic for the same types of mesoderm-derived cells, both basic and acidic FGF also stimulate new capillary growth (angiogenesis) in the chick chorioallantoic membrane and the rabbit cornea (*3, 5, 6*). Although the precise physiological role of the FGF's has not been established, they could promote the angiogenesis that accompanies the development of the corpus luteum, placenta, and fetus and may be involved in tissue repair. Also, the probable identity of basic FGF with tumor angiogenesis factor (*3*) implies an involvement with the neovascularization associated with the rapid growth of solid tumors.

The complete amino acid sequences of the bovine basic (*7*) and acidic (*8*) FGF's have been established. Basic FGF has 146 amino acids, although a form lacking the amino-terminal 15 amino acids has full mitogenic activity (*2*). Acidic FGF has 140 amino acids, and a form lacking the amino-terminal 6 amino acids is also biologically active. The two FGF's have a 55 percent amino acid sequence homology; slight homologies have also been detected between the FGF's and interleukin-1, and between a region of acidic FGF and neuropeptides such as neuromedin C, bombesin, substance P, and substance K (*6, 8*).

To clone the genes encoding acidic and basic FGF, we initially isolated a DNA fragment encoding part of the acidic FGF protein. Unique sequence oligonucleotide hybridization probes corresponding to amino acids 1 to 16 and 18 to 34 of bovine acidic FGF (*8*) were designed (Fig. 1), taking into account (i) the codons found in other sequenced bovine genes and (ii) the underrepresentation of the dinucleotide CpG in mammalian DNA (*9*). A bovine genomic library was screened in duplicate with these probes, and phages that hybridized to both probes were rescreened with a short, degenerate oligonucleotide probe corresponding to amino acids 7 to 12. Two phages, λBA2 and λBA3, hybridized to all three probes; these phages were purified and shown by restriction enzyme analysis to contain overlapping genomic sequences (*10*). A 250–base pair (bp) Alu I fragment was subcloned from λBA2, and its nucleotide sequence was determined (Fig. 1). An open translational reading frame, encoding the first 41 amino acids of bovine acidic FGF, could be identified within this sequence. Codon 42 is interrupted by an intron (*10*).

J. A. Abraham, A. Mergia, J. L. Whang, A. Tumolo, J. Friedman, K. A. Hjerrild, J. C. Fiddes, California Biotechnology, Inc., 2450 Bayshore Parkway, Mountain View, CA 94043.
D. Gospodarowicz, School of Medicine, Cancer Research Institute, University of California, San Francisco, CA 94143.