High molecular mass forms of basic fibroblast growth factor are initiated by alternative CUG codons

(hepatoma/protein sequencing/cDNA cloning/in vitro translation)

HERVÉ PRATS*,†, MOURAD KAGHAD‡, ANNE CATHERINE PRATS*, MICHAEL KLAGSBRUN§, JEAN MICHEL LÉLIAS‡, PHILIPPE LIAUZUN‡, PASCALE CHALON‡, JEAN PIERRE TAUBER†, FRANCOIS AMALRIC*, JOHN A. SMITH†, and DANIEL CAPUT‡

*Centre de Recherche de Biochimie et de Génétique Cellulaires, Centre National de la Recherche Scientifique, 118 route de Narbonne, 31062 Toulouse, France; †Unité 168, Institut National de la Santé et de la Recherche Médicale, Ranguel, Chemin du Vallon, 31054 Toulouse, France; §Laboratoire de Biologie Moléculaire, Sanofi-Elf Bio Recherches Boîte Postale 137, 31328 Labège cedex, France; ‡Departments of Surgery and Biological Chemistry, Children’s Hospital, Harvard Medical School, Boston, MA 02115; and †Departments of Molecular Biology and Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114

Communicated by Ernest Beutler, December 19, 1988

ABSTRACT A 6.75-kilobase human hepatoma-derived basic fibroblast growth factor (bFGF) cDNA was cloned and sequenced. An amino-terminus sequence generated from a purified hepatoma bFGF was found to correspond to the nucleotide sequence and to begin 8 amino acids upstream from the putative methionine start codon thought to initiate a 154-amino acid bFGF translation product. This sequence suggests that a form of bFGF of at least 163 amino acids exists. The hepatoma cDNA was transcribed in vitro into RNA; in vitro translation of this RNA generated three forms of bFGF with molecular masses of 18, 21, and 22.5 kDa. By use of an in vitro mutagenesis, it was found that the 22.5-kDa bFGF and possibly the 21-kDa form were initiated with CUG start codons. The 18-kDa bFGF was initiated with an AUG codon. By transfecting into COS cells human hepatoma bFGF cDNA and a construct from which the AUG initiator was eliminated, it was found that the higher molecular mass forms of bFGF were as biologically active as the 18-kDa form.

Basic fibroblast growth factor (bFGF) is a potent mitogen that stimulates growth and differentiation of a broad spectrum of mesodermal and neuro-ectodermal cell types (1). The angiogenic activity of bFGF is related to the mitogenicity of this factor towards vascular endothelial cells (2, 3). bFGF was originally purified from brain and pituitary as a 16- to 18-kDa monomeric polypeptide (4), but recently higher molecular mass bFGF forms of 25 kDa have been described (5). A primary translation product of 154 amino acids has been predicted from the nucleotide sequence of bFGF cDNA (6). However, a 157-amino acid placental bFGF has been described that includes at the amino-terminal end the methionine putatively claimed to be the bFGF start codon and 2 additional amino acids (7). In addition, we report here that hepatoma cells produce a 163-amino acid form of bFGF that extends even beyond the amino terminus of the placental bFGF. These results are surprising because no AUG codon in frame has been found upstream of the putative start codon that is at the beginning of the bFGF open reading frame (ORF). However, we have cloned a hepatoma bFGF cDNA larger than previously reported (6, 8). The 5'-end sequence of the hepatoma bFGF cDNA extends 466 nucleotides upstream of the putative methionine start codon. This leader sequence does not contain any AUG codons in frame with the bFGF nucleotide sequence but does contain several putative unusual start codons. Recently, the presence of an alternative CUG start codon able to initiate c-myc-1 protein as well as a viral protein were reported (9, 10). In the article we report that a 22.5-kDa high molecular mass form of hepatoma-derived bFGF is initiated with a CUG start codon.

MATERIALS AND METHODS

Cloning of the Human bFGF cDNA. Poly(A)+ RNA was prepared from human hepatoma cell line SK-HEP-1 as described by Cathala et al. (11). A cDNA library was constructed by using a pUC9 vector as described by Caput et al. (12) and consisted of about 106 recombinant bacterias. Synthetic oligonucleotide probes based on the coding sequence of the bFGF were used to screen the library by in situ colony hybridization.

Amino Acid Sequence Analysis. bFGF was purified from an extract of hepatoma cells as described (13) by a combination of Bio-Rex 70 (Bio-Rad) chromatography, heparin chromatography, and HPLC reverse-phase chromatography with a C3 column and was analyzed with an Applied Biosystems 470A sequencer (13).

DNA Sequence Analysis. The DNA sequencing was performed by the dideoxy chain-termination method (14) using restriction fragments cloned in M13mp18 and M13mp19 (Pharmacia).

In Vitro Mutagenesis and Plasmid Constructions. Point mutations on bFGF cDNA at positions 440 and 467 changing GGA to TGA and ATG to GTG, respectively, were generated by oligonucleotide-directed mutagenesis (15). Two other point mutations (positions 302, CTG → TTG or ATG) were introduced by replacing the Xho I–Not I fragment (positions 295–319) with a double-stranded synthetic oligonucleotide. For in vitro transcription, wild-type or point-mutated bFGF cDNAs, 3' truncated down to position 1711 (EcoRI), were subcloned into pSP64 (Promega Biotec) or pBluescript KS+ (Stratagene) vectors. For transfection experiments, the wild-type or point-mutated (ATG → GTG) bFGF cDNA up to position 945 (Bal I) were inserted into the pSVL vector (Pharmacia). A 5' deleted recombinant was constructed by inserting the fragment Apa I–Bal I (positions 455–943) into pSVL.

In Vitro Transcription and Translation. Linearized bFGF recombinant DNAs were transcribed by using SP6 (Promega Biotec) or T7 (Genoto, Geneva) RNA polymerase. After DNase I treatment, RNA was extracted twice with phenol/chloroform, precipitated with ethanol, and resuspended in...
were water. The RNAs, at a final concentration of 10–20 µg/ml, were translated in vitro at 37ºC for 1 hr with a rabbit reticulocyte lysate system (Promega Biotec) diluted twice in the presence of 0.5–1 mCi of [35]S-methionine (Amersham; 1 Ci = 37 GBq). Labeled proteins were analyzed by NaDodSO₄/12% PAGE followed by fluorography.

Transfection of DNA to COS Cells. Monkey COS cells (16) were grown in Dulbecco’s modified Eagle’s (DME) medium supplemented with 10% fetal calf serum (GIBCO/Bethesda Research Laboratories); 1.5–2 x 10⁶ cells were transfected with 20 µg of recombinant DNA in 5 ml of DME medium containing 10% (vol/vol) serum, 200 µg of DEAE-dextran per ml, and 100 µM chloramphenicol. After 5 hr at 37ºC, the transfection mixture was removed, and the cells were rinsed with phosphate-buffered saline (PBS) solution, treated for 1 min with 10% dimethyl sulfoxide in PBS, then rinsed twice with PBS, and incubated for 65 hr in DME medium containing 0.5% serum.

Mitogenic Activity of COS Cell Extracts. Cellular pellets of transfected COS cells were washed twice with PBS and resuspended at 10⁷ cells per ml in PBS supplemented with 2.0% bovine serum albumin. After sonication, the 12,000 x g pellets were discarded, and the mitogenic activity of the supernatants was determined by using adult bovine aortic endothelial (BAE) cells as follows: BAEB cells were seeded at 2 x 10⁴ cells per 35-mm dish in DME medium supplemented with 10% calf serum (Seromed). Various amounts of bFGF or COS cell extracts were added every 2 days. Mitogenic activity was determined by counting the BAE cells 5 days after seeding. The optimum activity was obtained when 2.5 µl of COS cell extract or 1 ng of Escherichia coli recombinant bFGF was used per ml of DME medium.

RESULTS

Nucleotide Sequence of a bFGF cDNA. A plasmid cDNA library was constructed from a human hepatoma cell line, SK-HEP-1, that expresses bFGF. cDNA clones encoding bFGF were isolated by hybridization with oligonucleotides. The clone designated pUC-SKI, containing the full-length representation of the major bFGF mRNA species, was sequenced (Fig. 1). The huge 3’ untranslated region of the cDNA (5823 nucleotides) is very A+T rich and contains eight potential polyadenylation sites. Some of these sites have been shown to generate bFGF RNAs with shorter 3’ untranslated sequences (6, 8). The sequence encoding the presumptive bFGF primary translation product starts at the

![Fig. 1. Nucleotide cDNA sequence (pUC-SKI) and predicted amino acid sequence of human bFGF. Potential polyadenylation sites are underlined. Arrows indicate the three potential initiation start sites discussed in this article.](image-url)
position 467 (with an AUG) and ends further downstream at the position 934. However, the bFGF reading frame extends roughly 300 nucleotides further upstream of the AUG start site at position 467 (Fig. 1). According to our nucleotide sequence, no other AUG is found in the upstream extension of the ORF. However, it is of interest that the bFGF nucleotide sequence contains four potential and unusual initiation codons [CUG at 302, 329, and 344; ACG (17, 18) at 407]. Our sequence is somewhat different in the G+C-rich region upstream of the AUG codon at 467 when compared with two previously published human bFGF cDNA nucleotide sequences. In these sequences, nucleotides missing at positions 463 (8) and 369 (6) modify the reading frame in such a way that the four unusual start codons are not in frame with the bFGF sequence. It is noteworthy that there is a high degree of homology between bFGF amino acid sequences (80%) deduced from the nucleotide sequences of human hepatoma-derived and bovine bFGF cDNAs upstream of the putative methionine start codon (Fig. 2).

N-Terminal Sequencing Analysis. Most of the bFGF purified from the hepatoma cell line was amino-terminally blocked (13). However, a small portion of the hepatoma bFGF sample had a free amino-terminal group available for Edman degradation, and the sequence of the 20-amino acid N-terminal protein fragment was obtained (Fig. 2). The hepatoma bFGF amino-terminal sequence was compared with the amino-terminal sequence of human placenta bFGF (7) and with the sequence of the primary translation products of bFGF predicted from both the nucleotide sequence of hepatoma bFGF cDNA and the nucleotide sequence of bovine bFGF cDNA (19). It is clear that the hepatoma bFGF fragment overlaps with the beginning of the bFGF ORF and extends 8 more amino acids upstream of the putative methionine start codon; 18 of the 20 amino acids in the hepatoma bFGF sequence are identical to those predicted from human hepatoma cDNA. Only the first 2 amino acids (−7 and −8) differ from the predicted sequence, but the identification of these 2 amino acids was tenuous. The first 2 amino acids (threonine at −1 and glycine at −2) of the hepatoma bFGF N-terminal extension corresponded to, and thus confirmed, the 2-amino acid N-terminal extension previously reported for placental bFGF (7) (Fig. 2). Thus, there is strong evidence that hepatoma cells synthesize a form of bFGF that is larger than that predicted by the putative ORF and that the identification of the actual human bFGF ORF (6) needs clarification.

In Vitro Translation of bFGF mRNA. A bFGF cDNA was cloned into an RNA transcription vector as described. The transcribed RNA was translated in vitro, and three protein bands of approximately equal intensity were identified (Fig. 3, lane 2). These three proteins had molecular masses of 18, 21, and 22.5 kDa. Replacement of ATG (at 467) by GTG abolishes synthesis of the 18-kDa but not the 21- or 22.5-kDa protein bands (Fig. 3, lane 3). Introduction of a TGA stop codon just upstream of ATG at 467 and in frame abolishes both the 21- and 22.5-kDa bFGF forms but not the 18-kDa form (Fig. 3, lane 4). Thus, 21- and 22.5-kDa bFGF forms must have other start codons upstream of the ATG codon at 467. The absence of any obvious AUG start codon upstream of the AUG at 467 prompted us to consider CUG and ACG as putative alternative start codons. From the hepatoma bFGF cDNA sequence, it appeared that CTG at 329, CTG at 342, and to a lesser extent ACG at 407 might be candidates to be initiator codons for the 21- and 22.5-kDa forms of bFGF. The Kozak context (20) suggested codons CTG at 302 and 344 to be particularly good candidates as initiator codons. Replacement of the CTG codon at 302 by TTG severely reduced but did not abolish the synthesis of the 22.5-kDa bFGF (Fig. 3, lane 5). On the other hand, replacement of CTG at 302 by ATG resulted in a dramatic 100-fold increase in synthesis of the 22.5-kDa bFGF (Fig. 3, lane 6).

In Vivo Expression of bFGF cDNA. To find out if the higher molecular mass forms of bFGF are biologically active, bFGF cDNA and bFGF cDNA with the ATG codon at 467 replaced by GTG were transfected into monkey COS cells. COS cell extracts were tested for the ability to stimulate ABAE cell proliferation (Fig. 4). Five microliters of extracts corresponding to 5 × 10⁴ cells transfected with both constructs stimulated ABAE cell proliferation equally and to the same extent as does 1 ng of recombinant bFGF per ml. Thus, initiation of bFGF at the putative ATG start codon at 467 is not required for synthesizing biologically active mitogen.

DISCUSSION

Analysis of bFGF structure (13, 21, 22) and the nucleotide sequence of bFGF cDNA (6, 8) has suggested that bFGF is a 154-amino acid protein with a 155-amino acid ORF initiated by an AUG start codon. However, a higher molecular mass form of bFGF extending beyond this AUG has been described in human placenta (7). In addition, an amino acid sequence of a bFGF fragment that we have isolated from human hepatoma cells (and that corresponds almost exactly to the sequence of bFGF predicted from human hepatoma bFGF cDNA) indicates that there must be at least a 163-amino acid form of bFGF beginning 8 amino acids upstream of the methionine start codon postulated to initiate 154-amino acid bFGF.

The existence of high molecular mass forms of bFGF, which include the methionine previously thought to be the initiator of bFGF translation, is surprising because analysis of the upstream nucleotide sequence fails to detect any other potential AUG initiation codons. One possible explanation is that there is 5′ sequence heterogeneity among bFGF mRNAs resulting from an alternative RNA splicing and leading to different translation products. A second possibility is that there are start codons other than AUG that can initiate bFGF synthesis. For example, CUG has been predicted to initiate c-myc-1 protein synthesis and has been demonstrated for a cell-surface antigen encoded by murine leukemia virus (9, 10). We have cloned a hepatoma bFGF cDNA with a larger 5′ extension and a slightly different sequence than has been reported previously (6, 8). The bFGF gene and the bFGF gene altered by in vitro mutagenesis have been translated in vitro and transfected into monkey COS cells. The results indicate that there are at least three possible initiation codons for bFGF, an AUG codon (at the position 467) that initiates the 18-kDa form of bFGF and two CUG codons (at positions
stream of the AUG start codon. The evidence that the CUG codon at 302 initiates synthesis of 22.5-kDa bFGF is that alteration of this codon to UUG virtually eliminates synthesis of the high molecular mass form of bFGF. There is also some circumstantial evidence that the CUG codon at 344 might initiate synthesis of 21-kDa bFGF because initiation at this site would give bFGF of the predicted size and because CUG at 344 interestingly has a repeat of the region around the codon CUG at 302 (GAGG/CTG/GGGG, Fig. 3B). Our in vitro translation studies indicate that initiation at unusual CUG start codons explains the existence of higher molecular mass forms of bFGF. However, hepatoma cells in culture appear to produce predominantly the 18-kDa form of bFGF (13). In fact, almost all cells and tissue produce 18-kDa or even lower molecular mass forms of bFGF. One likely explanation is that protease-mediated truncation of bFGF occurs either posttranslationally or during bFGF extraction. We have found (21) that hepatoma cells contain acid-activated proteases that cleave bFGF at the N-terminal end.

In addition, our experiments do not rule out the possibility that even higher molecular mass forms of bFGF exist, since it is not certain that the bFGF cDNA used to express bFGF is full length. Finally, although initiation at CUG codons is one possible mechanism for generating high molecular mass bFGF, other mechanisms such as alternative splicing might exist as well.

The efficiency of in vitro translation at the three initiation sites is low compared with that obtained when using globin mRNA. However, replacement of CUG at 302 with an AUG enhances by about 100-fold the synthesis of 22.5-kDa bFGF and also abolishes synthesis of the two lower mass forms of bFGF, suggesting (i) that CUG is less efficient than AUG as an initiation codon, as has been previously reported (23), and
(ii) that the low efficiency of the first start site scanned by ribosomes might favor the initiation of protein synthesis at other potential start sites, which is in agreement with the Kozak's model (24). Therefore, we hypothesize that unusual start codons might be a means for down-regulation of protein synthesis as well as an important feature of alternative initiation of protein synthesis.

The size and initiation site of bFGF do not seem to affect its biological activity. In transfection experiments, the 22.5-kDa bFGF initiated at the CUG codon at position 302 is as mitogenic as the 18-kDa bFGF initiated at the AUG codon at 467. Hepatoma cells produce mostly 18-kDa bFGF (13), but this molecular mass form could be a result of cell-mediated proteolysis of bFGF, which has been shown to occur in hepatoma cells (21). Higher molecular mass forms are more prominent in tissues such as brain (5) and rat cells (M. Klagsbrun, unpublished results). At present it is not clear how cells generate the various molecular mass forms of bFGF. There may be mechanisms that allow the cells to choose among various start codons and with various degrees of translation efficiency. In addition, bFGF size might be regulated by posttranslational steps that include protease-mediated processing. The relationship between the various forms of bFGF as well as the specific function of the bFGF amino-terminal sequences need clarification.

We are grateful to A. Minty, P. Ferrara, J. L. Darlix, A. Ezekowitz, and B. Beutler for comments and criticisms; C. Zanibellato and C. Minty for technical assistance; and F. Goguet-Chapuis and C. Kaghad for helping to edit the manuscript. This work was supported by Sanofi-Elf Bio Recherches and by the Association pour le Développement de la Recherche sur le Cancer (ARC); A.C.P. was the recipient of an ARC Fellowship.