
Human fibroblast interferon gene lacks introns

Richard M. Lawn, John Adelman, Arthur E. Franke, Catherine M. Houck, Mitchell Gross, Richard Najarian and David V. Goeddel

Department of Molecular Biology, Genentech, Inc., 460 Pt. San Bruno Blvd., South San Francisco, CA 94080, USA

Received 30 January 1981

ABSTRACT

A recombinant λ bacteriophage isolated from a human genome library contains the gene for fibroblast interferon (IFN- β 1). The DNA sequence of this gene is identical to the sequence of its mRNA and is devoid of introns.

INTRODUCTION

The past few years have seen a vast increase in our knowledge of the molecular biology of the interferon system. Several laboratories have isolated and sequenced cloned cDNA copies of human fibroblast interferon mRNA (IFN- β ; 1-3), and this mRNA has also been directly sequenced (4, 5). In contrast with the human leukocyte interferon (IFN- α) gene family which consists of at least eight distinct, but homologous genes (6-13), all of these laboratories (1-5) obtain only a single fibroblast interferon sequence, IFN- β 1. The amino acid sequence of IFN- β 1 deduced from the sequence of cloned cDNAs consists of a 21 amino acid signal peptide followed by the 166 amino acids which comprise mature fibroblast interferon. Evidence has been reported for the existence of a second type (IFN- β 2) of interferon mRNA in induced human fibroblasts (14). Cloned cDNA copies of IFN- β 2 mRNA, which bears only slight homology with IFN- β 1 mRNA, have recently been identified (M. Revel, personal communication).

We have previously shown that the IFN- β 1 gene and the entire leukocyte interferon gene family are located on human chromosome 9 (15). Furthermore, cloned chromosomal DNA fragments containing several LeIF genes have been characterized (11-13). The LeIF genes examined to date all lack intervening sequences. In this report we discuss the isolation and characterization of the human IFN- β 1 gene from a human gene library (16) and show that this gene also lacks intervening sequences.

MATERIALS AND METHODS

Phage Library Screening: A bacteriophage λ Charon 4A recombinant library constructed by Lawn *et al.* (16) containing Hae III and Alu I partial digestion fragments of the human genome was screened for IFN- β 1 containing clones by described methods (16, 17) with the exception that 10 percent dextran sulfate (Pharmacia) was added to the hybridization solution to accelerate hybridization (18). The radioactive probe employed was derived from fragments of the IFN- β 1 cDNA clone pFIF3 (3) labelled to $>10^8$ cpm/ μ g with 32 P (Amersham) using DNA polymerase (Large "Klenow" Fragment; New England Biolabs) extension of small, random calf thymus DNA primers (19). Phage DNA was purified as described (17).

Subcloning: The 1.6 kb Eco RI fragment of λ FIF1 which contained the IFN- β 1 gene was purified by electrophoresis in a 0.75 percent agarose slab gel (17), cut from the gel and electroeluted in dialysis tubing containing and submerged in a buffer of 9 mM Tris-borate (pH 8.3), 0.25 mM sodium EDTA at 90V for one hour. The eluted DNA fragment was purified by phenol and chloroform extractions, ethanol precipitated and subcloned into the Eco RI site of the plasmid pBR325 (20) or the replicative form of the phage mP7 (21), a derivative of m13. In addition a 2.1 kb Bgl II fragment of λ FIF1 containing the coding region was isolated as described above and subcloned into the Bam HI site of pBR322 (22) and of the phage mP7RF. Plasmid DNA was purified by the procedure of Blin and Stafford (23) and mP7 phage DNA as in (21).

DNA Sequencing: Plasmid DNA was cleaved with restriction endonucleases (New England Biolabs) as described by the manufacturer. Fragments were end labelled either at 5' termini (24) with (γ - 32 P)ATP (Amersham) and polynucleotide kinase (P-L Biochemicals) or at 3' termini with (α - 32 P)ATP and the large fragment of DNA polymerase. To accomplish 3' labelling, ~ 10 μ g of DNA was incubated with 0.1 mCi of (α - 32 P)dATP (2,000-3,000 Ci/mole; Amersham) and 200 μ M of the other three dNTPs in a 50 μ l volume of 6 mM Tris HCl (pH 8), 50 mM NaCl, 6 mM MgCl₂, 5 mM dithiothreitol, 100 μ g/ml BSA with 2 units of DNA polymerase large fragment for 60 min at 37°C. 32 P-labelled fragments were cleaved with a second restriction enzyme before isolation by gel electrophoresis and electroelution. Purified and labelled fragments were sequenced by the method of Maxam and Gilbert (24).

Dideoxy chain termination sequencing was performed on mP7 phage subclones as described previously (21, 25) using synthetic DNA primers

complementary to the border of the cloning vehicle (21) or a 14 nucleotide region of the known human FIF cDNA sequence (3). The tetradecanucleotide primer dCTGCTTCTTTGTAG was synthesized as described previously (26). Thin sequencing gels were run as described in (24). The use of 85 cm long 4 percent acrylamide gels allowed a maximum resolution of ~500 nucleotides per fragment. Gels were transferred to filter paper and dried before exposure to X-ray film (Kodak) at -70°C with or without DuPont Cronex intensifying screens.

Assay for interferon activity: Phage lysates were prepared for IF assay as follows. *E. coli* strain DP50 supF, grown to late log phase in NZYDT broth was infected with λ HFIF as previously described (27). Briefly, 100 μl of late log cells were infected with approximately 10^5 phage. After 10 min incubation at 37° in the presence of 10 mM MgCl_2 ; CaCl_2 , the cells were diluted into 5 ml NZYDT broth. Lysis occurred after 8-10 hrs at 37° . The cell debris was pelleted by centrifugation and 2.5 ml of the phage-containing supernatant was dialyzed overnight against PBS. The interferon activity in the supernatant was determined by comparison with NIH fibroblast interferon standard using the cytopathic effect (CPE) inhibition assay as described by Goeddel et al. (3).

RESULTS

The recombinant bacteriophage λ Charon 4A/human genome library constructed by Lawn et al. (16) was screened by hybridization (16, 17) with a radioactive probe derived from the human fibroblast interferon cDNA plasmid pFIF3 (3). A λ recombinant designated λ HFIF1 was obtained which contains a 7.4 kilobase (kb) segment of the human genome including the structural gene for fibroblast interferon. Digesting λ HFIF1 with the restriction endonuclease Eco RI produces fragments of 20 and 10.5 kb deriving from λ DNA, and fragments 2.8, 2.4, 1.6 and 0.6 kb from the human DNA insert. Only the 1.6 kb fragment hybridizes to the IFN- β 1 probe (data not shown). This 1.6 kb hybridizing Eco RI fragment was subcloned into the Eco RI site of pBR325 (20). A map of restriction endonuclease sites in the gene region and the sequencing strategy used are shown in Figure 1. DNA prepared from the pBR325 subclone was digested with various restriction endonucleases, end-labelled with ^{32}P and sequenced by the method of Maxam and Gilbert (24). A 2.1 kb Bgl II fragment containing the interferon gene was cloned into the single

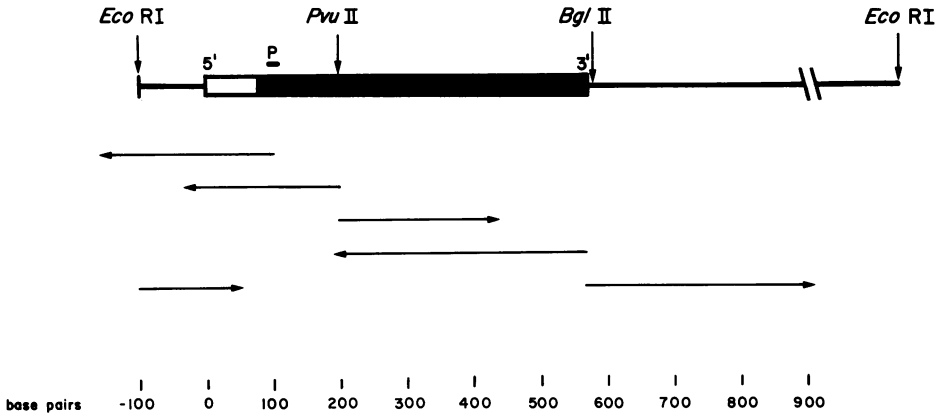


Figure 1. Sequencing strategy for fibroblast interferon gene. The region of λ HFIF1 DNA containing the IFN- β 1 gene is shown with the coding region for the mature peptide indicated by a solid box and the coding region of the signal peptide by an open box. Horizontal arrows represent DNA sequencing reactions initiating at the restriction endonuclease sites shown or at the site of a synthetic oligonucleotide indicated by P. This primer oligonucleotide, CTGCTTCTTTGTAG is complementary to nucleotides 88-101 of the gene sequence above.

stranded phage mP7 and sequenced by the dideoxy chain termination method (21, 25) employing oligonucleotide primers. The DNA sequence of the IFN- β 1 gene and flanking regions is presented in Figure 2.

The DNA sequence of the IFN- β 1 gene corresponds exactly to the published sequences of IFN- β 1 cDNA clones (1-3) and mRNA (4, 5). There are no introns in this IFN- β 1 gene. The 5' flanking region of the FIF gene in this recombinant terminates at the *Eco* RI site 102 nucleotides from the initiation codon. This *Eco* RI site is contained in the dodecameric *Eco* RI linker sequence which was used in the construction of the phage library (16). The sequence to the left of the *Eco* RI site is that of the Charon 4A vector and is not shown in Figure 2.

The 3' flanking sequence contains the hexanucleotide AATAAA which precedes by ~20 nucleotides the polyadenylation site of most eukaryotic mRNAs (28). The polyadenylation site of the FIF mRNA is located either 13 (3) or 20 (1, 2) nucleotides beyond this AATAAA sequence.

Lysates of *E. coli* infected with the recombinant phage λ FIF1 were examined for interferon activity as described in Materials and Methods. Approximately 20 units of interferon were obtained per ml of extract.

▼

GGCCATACCCTGGAGAAAGGACATTCTAACTGCAACCTTTCGAAGCCTTTGCTCTGGCACAAACAGGTAGTAGGGCGACACTGTTCTGGTTGTCAAC

met thr asn lys cys leu leu gln ile ala leu leu leu cys phe ser thr thr ala leu ser met ser tyr asn
 ATG ACC AAC AAG TGT CTC CTC CAA ATT GCT CTC CTG TTG TGC TTC ACT ACA GCT CTT TCC ATG AGC TAC AAC
 1 50

leu leu gly phe leu gln arg ser ser asn phe gln cys gln lys leu leu trp gln leu asn gly arg leu glu
 TTG CTT GGA TTC CTA CAA AGA AGC AGC AAT TTT CAG TGT CAG AAG CTC CTG TGG CAA TTG AAT GGG AGG CTT GAA
 100 150

tyr cys leu lys asp arg met asn phe asp ile pro glu glu ile lys gln leu gln phe gln lys glu asp
 TAC TGC CTC AAG GAC AGG ATG AAC TTT GAC ATC CCT GAG GAG ATT AAG CAG CTG CAG CAG TTC CAG AAG GAG GAC
 200

ala ala leu thr ile tyr glu met leu gln asn ile phe ala ile phe arg gln asp ser ser ser thr gly trp
 GCC GCA TTG ACC ATC TAT GAG ATG CTC CAG AAC ATC TTT GCT ATT TTC AGA CAA GAT TCA TCT AGC ACT GGC TGG
 250 300

asn glu thr ile val glu asn leu leu ala asn val tyr his gln ile asn his leu lys thr val leu glu glu
 AAT GAG ACT ATT GTT GAG AAC CTC CTG GCT AAT GTC TAT CAT CAG ATA AAC CAT CTG AAG ACA GTC CTG GAA GAA
 350

lys leu glu lys glu asp phe thr arg gly lys leu met ser ser leu his leu lys arg tyr tyr gly arg ile
 AAA CTG GAG AAA GAA GAT TTC ACC AGG GGA AAA CTC ATG AGC AGT CTG CAC CTG AAA AGA TAT TAT GGG AGG ATT
 400 450

leu his tyr leu lys ala lys glu tyr ser his cys ala trp thr ile val arg val glu ile leu arg asn phe
 CTG CAT TAC CTG AAG GCC AAG GAG TAC AGT CAC TGT GCC TGG ACC ATA GTC AGA GTG GAA ATC CTA AGG AAC TTT
 500

tyr phe ile asn arg leu thr gly tyr leu arg asn end
 TAC TTC ATT AAC AGA CTT ACA GGT TAC CTC CGA AAC TGA AGATCTCTAGCCTGTGCCTCTGGGACTGGACAATTGCTTCAAGCAT
 550 600

TCCTCAACCAGCAGATGCTGTTTAAGTGACTGATGGCTAATGACTGCATATGAAAGGACACTAGAAGATTTGAAATTTTATTAAATTATGAGTTAT
 650 700

TTTATTATTATTAATTTTATTTTGGAAATATAATTTTGGTGC[▼]AAAGTCAACATGGCAGTTTTAATTTGCAATTGATTATATAACCATCCATA
 750 800

TTATAAAATGGCAGTACCTATTAGTTGTTCTTTTTAAATATACCTGCAAAGTAGTATACTTTGGTTCTGCTTAAGGAATTTAAATTC
 850 900

Figure 2. Sequence of the fibroblast interferon gene. The DNA sequence of the coding strand of the IFN- β 1 gene and flanking regions is shown above. Nucleotides are numbered from the initiation codon of the signal peptide. Predicted amino acids are written above the nucleotide sequence in the protein coding region. The initial codon of the mature protein is enclosed in a box. Arrows indicate the 5' end (4, 5) and 3' end (1-3) of IFN- β 1 mRNA as reported by others. The common hexanucleotide preceding eukaryotic mRNA sites of polyadenylation (28) is underlined.

This synthesis is presumably due to transcription readthrough from a λ promoter, followed by translation of fibroblast interferon beginning either with the methionine AUG codon of the signal peptide or the AUG codon of the mature molecule.

DISCUSSION

In contrast to the human leukocyte interferon genes which number at least eight, the fibroblast gene we have analyzed may represent a single species. Southern blot hybridizations of total human DNA cleaved separately with several restriction endonucleases show only one fragment which hybridizes to an IFN- β 1 cDNA probe (8). Similar experiments with IFN- α cDNA probes yield multiple hybridizing fragments (8) reflecting the multiplicity of leukocyte interferon genes which has been demonstrated by sequence analysis of at least eight distinct cloned cDNA species. Recently, a distinct mRNA species with interferon-like activity has been isolated from human fibroblasts (14). This mRNA has less than 50 percent homology with the known IFN- β 1 species (M. Revel, personal communication). Therefore, the IFN- β 1 gene we describe may have more homology with the IFN- α gene family than with IFN- β 2, and an IFN- β 1 cDNA probe would not be expected to detect IFN- β 2 sequences in Southern hybridizations to total human genome DNA.

The directly determined NH₂-terminal amino acid sequence of human fibroblast interferon (29, 30) is identical to that predicted from the DNA sequence of the IFN- β 1 gene. Hence, it is likely that the IFN- β 1 gene contained in λ HFIF1 is the single human gene coding for the well studied interferon produced by stimulated fibroblast cells. Furthermore, the sequence of the gene contained in λ HFIF1 is identical with mRNA (4, 5) and cDNA sequences (1-3) of this gene. Hence this active gene is devoid of introns. The only other vertebrate nuclear genes described thus far which lack introns are five human IFN- α 1 genes (11-13). Yeast, sea urchin and drosophila histone genes and several other viral encoded and mitochondrial genes also lack introns (31-35).

The 3' flanking region of the IFN- β 1 gene contains the common AATAAA hexanucleotide (28) preceding the site of polyadenylation. Other features of the coding and flanking regions of interferon genes may be evaluated for functional significance as more interferons are characterized at DNA and protein levels. Since the interferons represent a multigene family it will be possible to study the functional relationship between related genes and proteins. The various interferon proteins may respond differentially in certain cell types to various inducers and have different target cell specificities. Eight different human leukocyte cDNA clones have been identified (8) and six of these have been efficiently expressed in E. coli (7, 9, D. Leung, P. Gray, D.

Goeddel; unpublished results). These interferons and bacterially produced IFN- β 1 (3) display varied activities in different animals and on cultured cells derived from various organs (P. Weck, N. Stebbing; personal communication).

The interferon gene family appears to be genetically linked. The fibroblast interferon gene (IFN- β 1) and the approximately ten leukocyte interferon (IFN- α) genes detectable by hybridization with cloned cDNA probes have all been localized to human chromosome 9 by blot hybridization to DNA derived from mouse-human hybrid cell lines (15). The close linkage of several of these genes has been shown by the characterization of λ recombinant clones containing two separate IFN- α genes each (11-13). It remains to be seen whether all members of the human interferon gene family are clustered, as the relationship of the presently characterized interferon genes to immune interferon (IFN- γ) and the newly reported fibroblast type (IFN- β 2) interferon has yet to be elucidated.

ACKNOWLEDGEMENT

We thank Roberto Crea for supplying the deoxyoligonucleotides used for DNA sequencing.

REFERENCES

1. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. and Muramatsu, M. (1980) Gene 10, 11-15.
2. Derynck, R., Content, J., DeClercq, E., Volckaert, G., Tavernier, J., Devos, R., and Fiers, W. (1980) Nature 285, 542-547.
3. Goeddel, D.V., Shepard, H.M., Yelverton, E., Leung, D. and Crea, R., Sloma, A., and Pestka, S. (1980) Nucl. Acids Res. 8, 4057-4074.
4. Houghton, M., Stewart, A.G., Doel, S.M., Emtage, J.S., Eaton, M.A.W., Smith, J.C., Patel, T.P., Lewis, H.M., Porter, A.G., Birch, J.R., Cartwright, T. and Carey, N.H. (1980) Nucleic Acids Res. 8, 1913-1931.
5. Houghton, M., Eaton, M.A.W., Stewart, A.G., Smith, J.C., Doel, S.M., Catlin, G.H., Lewis, H.M., Patel, T.P., Emtage, J.S., Carey, N.H. and Porter, A.G. (1980) Nucleic Acids Res. 8, 2885-2895.
6. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Escodi, J., Boll, W., Cantell, K. and Weissmann, C. (1980) Nature 284, 316-320.
7. Goeddel, D.V., Yelverton, E., Ullrich, A., Heyneker, H.L., Miozzari, G., Holmes, W., Seeburg, P.H., Dull, T., May, L., Stebbing, N., Crea, R., Maeda, S., McCandliss, R., Sloma, A., Tabor, J.M., Gross, M., Familletti, P.C. and Pestka, S. (1980) Nature 287, 411-416.
8. Goeddel, D.V., Leung, D.W., Dull, T.J., Gross, M., Lawn, R.M., McCandliss, R., Seeburg, P.H., Ullrich, A., Yelverton, E. and Gray, P.W. (1981) Nature (in the press).
9. Yelverton, E., Leung, D., Weck, P., Gray, P.W. and Goeddel, D.V. (1981) Nucleic Acids Res. 9 (in the press).

10. Streuli, M., Nagata, S. and Weissmann, C. (1980) *Science* 209, 1343-1347.
11. Nagata, S., Mantei, N. and Weissmann, C. (1980) *Nature* 287, 401-408.
12. Lawn, R.M., Adelman, J., Dull, T.J., Gross, M., Goeddel, D.V. and Ullrich, A. (1981) *Science* (in the press).
13. Ullrich, A., Dull, T., Gray, A., Najarian, R., Goeddel, D.V., Seeburg, P. and Lawn, R.M. (1981) in preparation.
14. Sehgal, P.B. and Sagar, A.U. (1980) *Nature* 288, 95-97.
15. Owerbach, D., Rutter, W.J., Shows, T.B., Gray, P., Goeddel, D.V. and Lawn, R.M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* (in the press).
16. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) *Cell* 15, 1157-1174.
17. Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) *Cell* 15, 687-701.
18. Wahl, G.W., Stern, M. and Stark, G.R. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 3683-3687.
19. Taylor, J.M., Illemensee, R. and Summer, S. (1976) *Biochim. Biophys. Acta* 442, 324-330.
20. Bolivar, F. (1978) *Gene* 4, 121-134.
21. Messing, J., Crea, R. and Seeburg, P.H. (1981) *Nucleic Acids Res.* (in the press).
22. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) *Gene* 2, 95-113.
23. Blin, N. and Stafford, D.W. (1976) *Nucleic Acids Res.* 3, 2303-2308.
24. Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
25. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.
26. Crea, R. and Horn, T. (1980) *Nucleic Acids Res.* 8, 2331-2348.
27. Blattner, F.R., Williams, B.G., Blechl, A.E., Thompson, K.D., Faber, H.E., Furlong, L.A., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977) *Science* 196, 161-169.
28. Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature* 263: 211-214.
29. Knight, E. Jr., Hunkapillar, M.W., Korant, B.D., Hardy, R.W.F. and Hood, L.E. (1980) *Science* 207, 525-526.
30. Stein, S., Kenny, C., Friesen, H-J., Shively, J., Del Valle, V. and Pestka, S. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77, 5716-5719.
31. McKnight, S. (1981) *Nucleic Acids Res.* (in the press).
32. Smith, M., Leung, D.W., Gillam, S., Astell, C.R., Montgomery, D.L. and Hall, B.D. (1979) *Cell* 16, 753-761.
33. Sures, I., Lowry, J., Kedes, L. (1978) *Cell* 15, 1033-1044.
34. Holland, J.P. and Holland, M.J. (1980) *J. Biol. Chem* 255, 2596-2605.
35. Schaffner, W., Kunz, G., Deatwyler, H., Telford, J., Smith, H.O. and Birnstiel, M.L. (1978) *Cell* 14, 655-671.