Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5305–5309, September 1981
Biochemistry

Structure of a chromosomal gene for human interferon β

(genomic clone/DNA sequence/nuclease S1 mapping/sequence homology)

SHIGEO OHNO AND TADATSUGU TANIGUCHI

Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Toshima-ku, Tokyo 170, Japan

Communicated by Igor Tamm, March 9, 1981

ABSTRACT **We have cloned and analyzed a chromosomal DNA segment containing the human interferon β_1 gene from a human gene library. The nucleotide sequence of the protein-coding and the noncoding regions of the chromosomal gene was identical to the cDNA sequence reported previously. In the region upstream from the putative transcription initiation site, significant nucleotide sequence homology was observed between interferon β_1 and α_1 genes. This region thus may play a role in expression of the interferon genes. From the sequence data and the result of nuclease S1 mapping experiments, we conclude that, like the interferon α_1 gene, the interferon β_1 gene is devoid of intervening sequences.**

Much information about genes for human interferon α (IFN-α) as well as interferon β (IFN-β) has been accumulated in the past few years. We have previously reported the cloning of the cDNA for human IFN-β_1,* an IFN produced by human fibroblasts in response to poly(I)·poly(C) (1, 5). Nucleotide sequence analysis revealed that IFN-β_1 consists of 166 amino acids and arises from a precursor containing 21 additional amino acids (6–9).

In addition to IFN-β_1 cDNA, two of the cDNAs for IFN-α have been cloned and their sequences have been determined (10–12). The cloned cDNAs also directed the synthesis of active IFN molecules in *Escherichia coli* (10, 13–15). Comparison of the cDNA sequence of IFN-α_1 and β_1 showed apparent homology both in amino acid sequence and in nucleotide sequence, and we thus concluded that the two genes were derived from a common ancestor (16).

Analysis of the human chromosomal DNA revealed the existence of at least eight distinct IFN-α structural genes (17). So far, little is known about chromosomal gene(s) for IFN-β. It has been reported that human fibroblasts produce an additional IFN, IFN-β_2, whose mRNA size is distinct from that of IFN-β_1 (3, 4).

In order to study structural organization as well as the mechanism of the expression of human IFN-β genes, it was desirable to obtain a chromosomal DNA segment containing a human IFN-β structural gene and its regulatory region. In addition, it was of great interest to compare chromosomal genes for IFN-β with those of IFN-α in view of the close relationship between them (16).

In this article we report cloning and analysis of the chromosomal gene for IFN-β_1. We show that, like IFN-α_1, the chromosomal IFN-β_1 gene lacks intervening sequences. We also show the occurrence of significant nucleotide sequence conservation between IFN-α_1 and IFN-β_1 gene upstream from the putative initiation site for transcription.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Preparation of DNA. Human chromosomal DNA was prepared from the nucleus fraction of human foreskin fibroblast strain DIP2, kindly provided by S. Kobayashi (5), essentially according to the published procedure (18). Plasmid DNA and phage DNA were prepared by the published procedures (19, 20).

Isolation of IFN-β Genomic Clones. Two sets of a human gene library (21) were generously provided by T. Maniatis. One contained fragments of fetal human chromosomal DNA, generated by partial cleavage with *Hae* III and *Alu* I restriction endonucleases and joined with *Eco*RI linkers to λ Charon 4A arms. Another one was prepared in a similar way except that the human DNA was partially cleaved with *Eco*RI and joined to the arms. IFN-β_1-specific clones were screened by the *in situ* procedure of Benton and Davis (22), using as a probe the 0.6-kilobase (kb) *Hin*cII/*Bgl* II fragment of TpIF319-13 cDNA (1) (cDNA insert). A hybrid phage clone termed λHIFNβ-121 was characterized as described later.

The chromosomal DNA fragment containing the IFN-β_1 gene was inserted into plasmid pBR322 as follows: DNA from hybrid phage λHIFNβ-121 was first digested with *Eco*RI. Cohesive ends were rendered flush by DNA polymerase Klenow fragment, tailed with dCMP residues by terminal transferase, and inserted into the *Pst* I site of pBR322 as described (23). This hybrid plasmid has the original *Eco*RI sites of the chromosomal DNA restored in both ends as well as *Pst* I sites of the pBR322 in both ends. A hybrid plasmid containing a 1.8-kb genomic fragment, termed pHIFNβ-121-312, was further characterized.

Filter Hybridization. Nick-translated ^{32}P-labeled DNA probes were prepared essentially according to the procedure of Roop *et al.* (24). Blotting analysis of the DNA transferred to nitrocellulose filters was carried out according to the procedure described by Kataoka *et al.* (25), which is based on the method of Southern (26).

Restriction Enzyme Cleavage Site Mapping and DNA Sequence Analysis. The maps of restriction sites were constructed by the procedure of Smith and Birnstiel (27), and confirmed and refined by the DNA sequence analysis. The nucleotide sequence was determined by the Maxam–Gilbert procedure (28).

S1 Nuclease Protection Mapping. Mapping of the 5'-terminus of IFN-β_1-mRNA was conducted by following the published procedure (29, 30).

Recombinant DNA Safety Procedures. The construction, screening, and propagation of recombinant bacteriophage and

Abbreviations: IFN, interferon; kb, kilobase(s).

* The interferon coded by TpIF319-13 cDNA (1) corresponds to that characterized by Knight *et al.* (2) and is termed IFN-β_1 here because of the possible existence of another β-type interferon (3, 4).

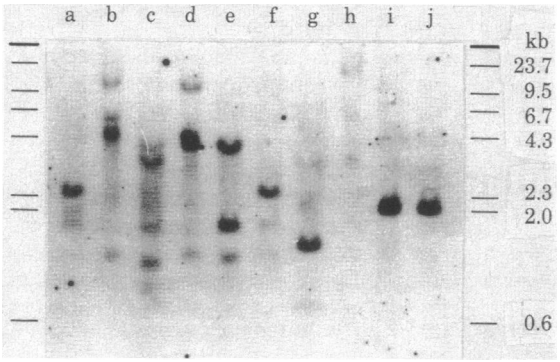

FIG. 1. Blot hybridization analysis of human chromosomal DNA. Chromosomal DNA from human fibroblasts was digested with restriction enzymes as indicated below, and resulting fragments were separated by 0.8% agarose gel electrophoresis. They were then transferred to a nitrocellulose filter and hybridized to the nick-translated TpIF319-13 DNA (1.2×10^7 cpm, specific activity 2.7×10^8 cpm/μg). Hybridization was carried out for 20 hr at 65°C. The hybridization solution (1 ml) contained 1 M NaCl, 50 mM Tris·HCl (pH 7.4), 10 mM EDTA, 0.2% Ficoll, 0.2% polyvinylpyrrolidone, 0.2% bovine serum albumin, sonicated and denatured *E. coli* DNA at 50 μg/ml, 0.1% sodium dodecyl sulfate, and the nick-translated TpIF319-13 DNA. After hybridization the filter strip was washed with 15 mM NaCl/1.5 mM sodium citrate, pH 7.0, containing 0.1% sodium dodecyl sulfate at 65°C for a total of 60 min with two changes of the washing solution (26). The lanes contain the DNA digested by the following restriction enzymes. a, *Pst* I; b, *Hin*dIII, c, *Hin*dIII and *Pst* I; d, *Hin*dIII and *Bam*HI; e, *Hin*dIII and *Eco*RI; f, *Pst* I and *Bam*HI; g, *Pst* I and *Eco*RI; h, *Bam*HI; i, *Eco*RI; j, *Bam*HI and *Eco*RI. Numbers on the right indicate the positions of size markers.

plasmid were conducted in a P-3 laboratory at the Cancer Institute in accordance with the guidelines for research involving recombinant DNA molecules issued in March 1979 by the Ministry of Education, Science and Culture of Japan.

RESULTS

Blotting Analysis of the Chromosomal DNA. To study organization and structure of the human IFN-β_1 gene, we first carried out Southern blot analysis of the chromosomal DNA isolated from human fibroblasts, and a typical result is shown in Fig. 1. With TpIF319-13 cDNA as the probe, a *Pst* I digest gave rise to two main positive bands (2.4 and 1.6 kb), whereas *Eco*RI digestion of the DNA gave a single major band at around 2 kb as well as two minor bands at 6 and 15 kb. On the other hand, both *Hin*dIII and *Bam*HI digests gave three or four positive bands (*Hin*dIII: 11, 6, 4.5, and 1.2 kb; *Bam*HI: 15, 6, and 3.5 kb). Moreover, *Eco*RI/*Hin*dIII double digests also gave three positive bands of 4.0, 1.7, and 1.2 kb. The results thus indicate that the chromosomal DNA contains sequences (or genes) related to IFN-β_1 gene or that IFN-β_1 gene is split by several intervening sequences, or both. Alternatively, those fragments may come from the region not specific for the IFN-β gene (see *Discussion*).

Isolation of the Recombinant Phage Containing IFN-β_1 Gene. In order to characterize further the IFN-β_1 gene as well as its related sequences, we next decided to clone the DNA segments containing the above sequences. The human gene bank prepared by Lawn *et al.* (21) was screened by an *in situ* procedure (22), using the ^{32}P-labeled TpIF319-13 cDNA insert as the probe. Eleven hybridization-positive phage clones were isolated from approximately 1,000,000 plaques from the *Hae* III/*Alu* I library. When these hybrid phage DNAs were digested with *Eco*RI and analyzed by blot hybridization, all DNA preparations gave a single hybridization-positive band of 1.8 or 1.5 kb (result not shown). The *Eco*RI restriction analysis of these 11 phage clone DNAs indicated that these DNA fragments originated from the same region of chromosomal DNA. This region contains a 1.8-kb *Eco*RI fragment that apparently corresponds to the major *Eco*RI fragment detected in the chromosomal DNA (The 1.5-kb fragment did not extend to the original *EcoRI* site that is upstream from the IFN-β_1 gene) (see Fig. 1, lane i). On the other hand, by screening the same number of plaques, only one hybridization-positive phage clone was isolated from the *Eco*RI library. The phage DNA did not contain the 1.8-kb *Eco*RI fragment but contained a 6-kb *Eco*RI fragment that hybridized with the cDNA probe, and thus corresponds to the 6-kb *Eco*RI fragment detected in the chromosomal DNA.

We therefore analyzed extensively one of those clones, termed λHIFNβ-121, that contains the 1.8-kb *Eco*RI fragment. The restriction map of the chromosomal segment from λHIFNβ-121 is shown in Fig. 2*b*. The restriction analysis indicates that

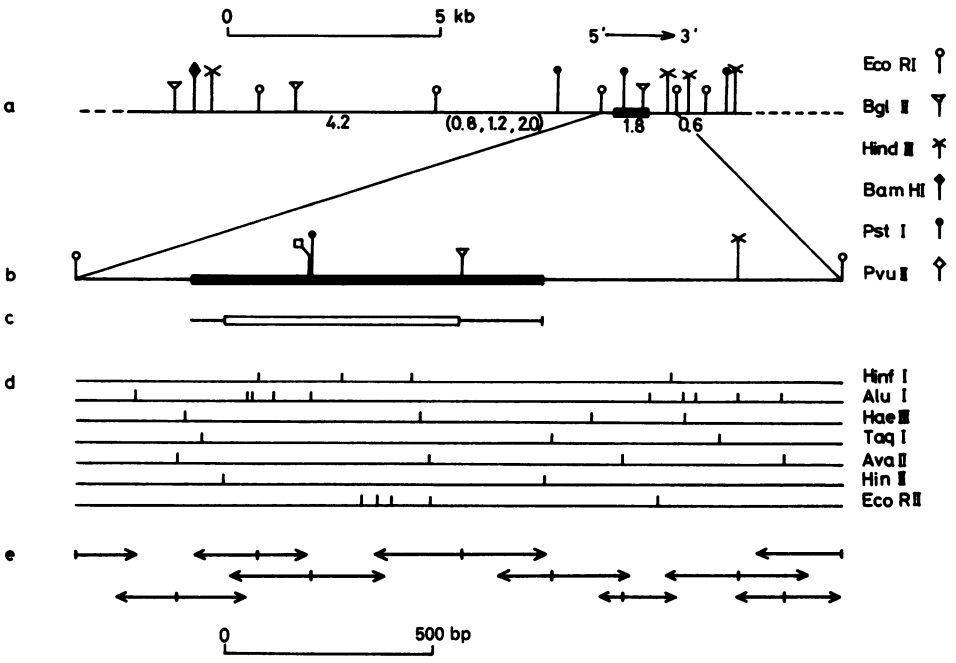

FIG. 2. Restriction enzyme cleavage map of human IFN-β_1 gene and sequencing strategy. (*a*) Restriction map of the 15-kb chromosomal DNA segment cloned in λHIFNβ-121. The map was constructed by blot hybridization analysis. The length of the fragments derived by *Eco*RI digestion are shown in kb. The order of the *Eco*RI fragments 0.8, 1.2, and 2 kb long is still to be determined. The broken line indicates the arms of the vector DNA from Charon 4A. (*b* and *d*) Detailed restriction map of the 1.8-kb *Eco*RI fragment. The restriction map was constructed by published procedures (27). It was confirmed and refined by the nucleotide sequence data as shown in Fig. 3. The black box shows the region from which the mRNA is transcribed. (*c*) Sequence arrangement of the IFN-β_1 mRNA. Open box indicates the protein coding region. (*e*) Strategy for sequence determination. Arrows indicate the direction and extent of sequencing of each fragment analyzed; the vertical bar at the end of each arrow represents ^{32}P-labeled 5'-terminal phosphate. bp, Base pairs.

besides the 1.8-kb *Eco*RI fragment this chromosomal segment confers the following hybridization-positive fragments obtained in blot analysis of the total chromosomal DNA: *Pst* I (2.4 and 1.6 kb); *Hin*dIII (11 kb); *Hin*dIII/*Pst* I (1.6 and 1.1 kb); *Hin*dIII/*Bam*HI (11 kb); *Hin*dIII/*Eco*RI (1.7 kb); *Pst* I/*Bam*HI (2.4 and 1.6 kb); *Pst* I/*Eco*RI (1.4 and 0.6 kb); and *Eco*RI/*Bam*HI (1.8 kb). However, other positive fragments obtained in Fig. 1 cannot be attributed to the cloned DNA (see *Discussion*). We next constructed a detailed restriction map of the 1.8-kb *Eco*RI fragment by inserting it into pBR322 DNA as shown in Fig. 2*c*.

As previously reported (6, 7) TpIF319-13 cDNA, contains the entire protein-coding region as well as the 3′-noncoding region of the mRNA but it lacks part of the 5′-noncoding region. The comparison of the restriction map of the genomic DNA segment with that of the cDNA indicates that the genomic DNA segment contains all the restriction sites from the *Taq* I site located upstream from the first ATG, to the *Bgl* II site, located downstream from the TAG termination sequence. The result therefore indicated that the chromosomal DNA segment contains the genomic sequence of IFN-β_1 and that IFN-β_1 gene contains no large intervening sequence, at least in its protein-coding region.

Nucleotide Sequence Analysis of the Chromosomal IFN-β_1 Gene. We next determined the nucleotide sequence of the 1.8-kb fragment in order to ascertain that the segment in fact has the sequence capable of coding for IFN-β_1 mRNA and in order to elucidate the structure surrounding the IFN-β_1 gene. As shown in Fig. 3, it is clear that the *Eco*RI 1.8-kb chromosomal DNA fragment contains the gene for IFN-β_1 mRNA, because the nucleotide sequences of both coding and noncoding regions are identical to those of the cDNA clone obtained by us and

```
                                                                                GAATTCTCAGGTCGTTTGCTTTCCTTTGCTTTCTCCCAAGTCTTGTTTTACAATTTG
                                                                                                                    -350

  CTTTAGTCATTCACTGAAACTTTAAAAAACATTAGAAAACCTCACAGTTTGTAAATCTTTTTCCCTATTATATATATCATAAGATAGGAGCTTAAATAAA
-300                                                                                                  -250

  GAGTTTTAGAAACTACTAAAATGTAAATGACATAGGAAAACTGAAAGGGAGAAGTGAAAGTGGGAAATTCCTCTGAATAGAGAGAGGACCATCTCATATA
-200                                                                                                  -150

  AATAGGCCATACCCACGGAGAAAGGACATTCTAACTGCAACCTTTCGAAGCCTTTGCTCTGGCACAACAGGTAGTAGGCGACACTGTTCGTGTTGTCAAC
-100                                                                                                   -50                                                                                                                -1

  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 TCC 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 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
  TAC TTC ATT AAC AGA CTT ACA GGT TAC CTC CGA AAC TGA AGATCTCCTAGCCTGTGCCTCTGGGACTGGACAATTGCTTCAAGCATT
                                      550                                                                                                  600

  CTTCAACCAGCAGATGCTGTTTAAGTGACTGATGGCTAATGTACTGCATATGAAAGGACACTAGAAGATTTTGAAATTTTTATTAAATTATGAGTTATTT
                                      650                                                                                                  700

  TTATTTATTTAAATTTTATTTTGGAAAATAAATTATTTTTGGTGCAAAAGTCAACATGGCAGTTTTAATTTCGATTTGATTTATATAACCATCCATATTA
                                      750                                                                                                  800

  TAAAATTGCCAAGTACCTATTAGTTGTTCTTTTTAAAATATACCTGCAAAGTAGTATACTTTCTGGCCCCTGCCTTTAAGGAATTTAAAATTCAAGAAAG
                                      850                                                                                                  900

  CCATGATGGAATATATAAGGTAAGAGACAATAAGGGGACCTGAACCTTATGGGGGAATAAATATGGCATGAACTGCTGTGGGATTAAAAGAGAAAAGGAA
                                      950                                                                                                 1000

  AGCTGGAGGGTCTGGAACTAAACCTGGGGTTCCCATTCCTCCTACTGTGTGTTCCAGATTCTCTCATCATAAAGTTAGAATTGAGCTGGCCATCAGGAAT
                                     1050                                                                                                 1100

  AGCCAGAGGAATATGTCAGCTTTTGTGTTCTCCCTAACCTTCCCCAGTTATTTGGGGGATCACTTTGCTCCTCGAAAGATTTTTAAATAATTATGTGCCC
                                     1150                                                                                                 1200

  CCCACCATCCCTGCAAGCTTAAGGGTGAGAAGTCCCATTTACTTCCATGACACTATTAAGCAGCAATCTCTTTATTCTGCTCATCATGGGACAGCCAAGA
                                     1250                                                                                                 1300

  TGTGTGGGTATCTTAGGGGAGCTGTGGGTCCCTGTCTGCTGGCATGGCACAGGCATCAGAGGAAGAAGAACCTTTTTATACCCTAGCCATCTGGTTAGTT
                                     1350                                                                                                 1400

  TTCTCCCTAGTTTTTCAAAAAACTAAGCCTGCTTCCAGTCCCCACTGCCTTGTTCATACAGAATTC
                                     1450
```

FIG. 3. Nucleotide sequence of the 1.8-kb *Eco*RI fragment containing the IFN-β_1 gene. The nucleotide sequence was determined according to the procedure of Maxam and Gilbert (28), following the strategy indicated in Fig. 2*e*. The amino acid sequence of IFN-β_1 (indicated by capital letters) and its putative signal peptide (indicated by lower-case letters) was deduced by comparing the nucleotide sequence with that of TpIF319-13 cDNA (6). The vertical arrows at −73 to −75 indicate the location of the probe termini obtained by S1 nuclease mapping analysis as shown in Fig. 5; the arrows at 764 to 766 indicate the position of poly(A) attachment.

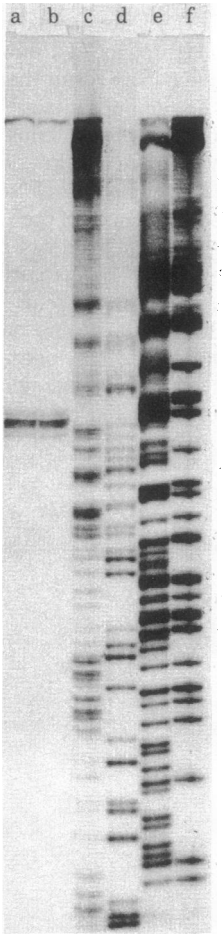

FIG. 4. S1 nuclease protection mapping of the 5′ terminus of IFN-β_1 mRNA. A 438-base pair *Eco*RI/*Hin*fI restriction fragment of the plasmid pHIFNβ121-312, which extends from −352 to +86, was prepared with a ^{32}P-labeled 5′ terminus at the *Hin*fI site. The fragment was denatured and hybridized to poly(A)-RNA from poly(I)·poly(C)-induced human fibroblasts (1) under the conditions described in refs. 29 and 30. After hybridization, the mixture was incubated with S1 nuclease and analyzed by 8% polyacrylamide gel electrophoresed (28). To provide chain length markers, the relevant DNA fragment was degraded by the Maxam–Gilbert procedure (28) and electrophoresed in parallel. Lane a, the ^{32}P-labeled DNA fragment (0.1 pmol, 35,000 cpm) was hybridized to 2 μg of induced fibroblast poly(A)-RNA and digested with 30 units of nuclease S1. Lane b, as a, 70 units of nuclease S1 was used. Lanes c, d, e, and f contain A+G, G, T+C, and C degradation products, respectively.

others (6–9). In addition, from the sequence analysis, there is no evidence for the presence of intervening sequences in the IFN-β_1 gene.

The sequence analysis upstream from the coding region revealed the presence of the T-A-T-A-A-A sequence, designated as the TATA box, which has been proposed as one of the consensus sequences of the genes from eukaryotes for the transcription initiation by RNA polymerase II (31).

Location the 5′ Terminus of the IFN-β_1 mRNA. In order to determine the site of the 5′ end of IFN-β_1 mRNA in the cloned chromosomal DNA, an S1 nuclease mapping experiment was carried out (29, 30). A 438-base pair *Eco*RI/*Hin*fI restriction fragment of the chromosomal DNA that extends from −352 to +86 was prepared with a ^{32}P-labeled 5′ terminus at the *Hin*fI cleavage site, hybridized with poly(A)-RNA from induced human fibroblasts, and digested with S1 nuclease. The sizes of the protected DNA fragment(s) were analyzed by polyacrylamide gel electrophoresis, using as a size marker the relevant DNA fragment degraded by the Maxam–Gilbert procedure (28).

As shown in Fig. 4, three major protected fragments and three minor ones were detected. The nucleotide lengths of these protected fragments are 161, 160, and 159 for the major ones and 163, 162, and 157 for the minor ones (32). The results thus indicate that the sequence for the 5′ terminus of the mature IFN-β_1 mRNA lies around nucleotide −73 to −75 of the DNA. This is to be compared with the nucleotide sequence of the IFN-β_1 cDNA, obtained by Houghton *et al.* (7), which extended up to T at position −72.

DISCUSSION

We have cloned and analyzed a chromosomal segment that contains the human IFN-β_1 gene. Complete nucleotide sequence analysis of the 1.8-kb *Eco*RI fragment revealed that the genomic DNA contained a sequence identical to the previously reported cDNA sequence in both the coding and the noncoding regions of IFN-β_1 mRNA. Taken together with the fact that the 1.8-kb fragment is predominant in the blot hybridization analysis (Fig. 1), it is most likely that the cloned DNA corresponds to the gene that gives rise to IFN-β_1 mRNA *in vivo*.

In addition to the IFN-β_1 gene we cloned, blot analysis of the total chromosomal DNA suggests the existence of IFN-β_1-related sequences. In fact, we also cloned a genomic DNA segment that, upon *Eco*RI digestion, gives rise to a 6-kb fragment instead of a 1.8-kb fragment during screening of the *Eco*RI gene library. On the other hand, Blattner *et al.* (33) reported the presence of A+T-rich DNA sequences in mammalian genomes that occasionally gave false positive clones upon screening gene libraries. Because our DNA probe used for blot analysis is TpIF319·13 plasmid DNA, it contains two stretches of A+T-rich tracts in both edges of the cDNA (5). At present, we cannot rule out the possibility that those DNA fragments that are not attributable to the cloned IFN-β_1-specific gene come from such a region. Although screening of the two gene libraries was carried out by using the *Hin*cII/Bgl II fragment of the cDNA as the nick-translated probe (see Fig. 3), it therefore remains to be seen whether the cloned 6-kb *Eco*RI fragment described above originates from any specific sequence for another β-type IFN gene(s) or from pseudogene(s) as reported in other genes (34).

From the nuclease S1 mapping experiments described above, we estimate the site in the genomic DNA corresponding to the 5′ terminus of the mature IFN-β_1 mRNA at 74 ± 2 nucleotides upstream from the ATG (Figs. 3 and 5). Because many if not all of the genes from eukaryotes contain a sequence, T-A-T-A-A-A that is also found in the IFN-β_1 gene (discussed below) approximately 25 nucleotides upstream from the putative transcription initiation site and because many of the eukaryotic pri-

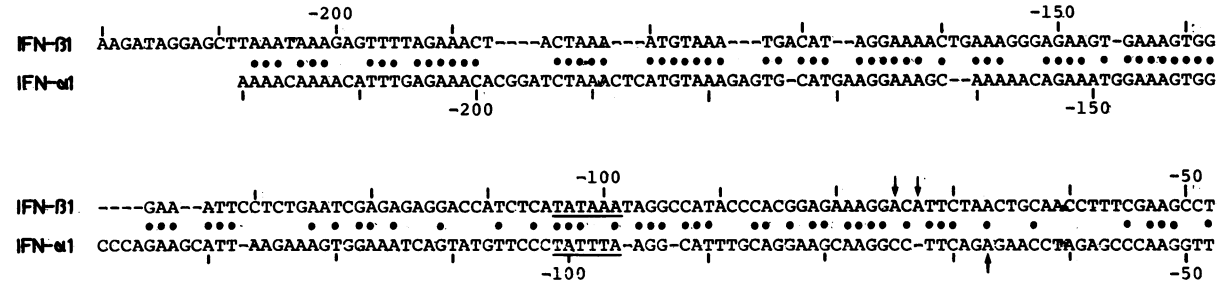

FIG. 5. Comparison of the nucleotide sequences upstream from the putative transcription initiation sites of the IFN-β_1 and IFN-α_1 genes. In aligning the sequences for IFN-β_1 and IFN-α_1, gaps were introduced to maximize homology. Dots indicate identical nucleotide sequences. Arrows indicate putative transcription initiation sites. The TATA box is underlined. Sequence data for the IFN-α_1 gene are from Nagata *et al.* (17).

mary transcripts seem to start with A (31), transcription of the human IFN-β_1 gene may start from A either at position -73 or at position -75.

From the nucleotide sequence analysis of the genomic DNA fragment as well as from the nuclease S1 mapping experiments, no evidence was obtained for the presence of any intervening sequences in the human IFN-β_1 gene. Interestingly, the same observation was reported with respect to human IFN-α genes (17). Although there are cases in which intervening sequences seem to play a role in efficient gene expression (35, 36), the absence of this sequence in both IFN-β_1 and IFN-α genes in addition to several protein-coding genes from eukaryotes (37, 38) further argues against the general role of this sequence in gene expression.

As shown in Fig. 3, the sequence T-A-T-A-A-A, designated as the TATA box (31), is also present in the IFN-β_1 gene. This sequence may play a role in correct initiation of the mRNA transcription; however, adherence to this consensus sequence is not absolute (31). In fact, the corresponding sequence for the IFN-α_1 gene seems to be rather divergent (17). It has been reported that, in addition to the TATA box, sequences upstream from the box are also important for efficient transcription (31, 39). Again, the nucleotide sequence of this region varies from gene to gene and the homologous sequence suggested by Corden *et al.* (31), -G-C-Y-C-A-A-T-C-C- (Y = pyrimidine), was found neither in the IFN-β_1 gene nor in the IFN-α_1 gene. From this point of view, it is worth noting that, upstream from the TATA box, extensive sequence homology is seen between the IFN-α_1 gene and the IFN-β_1 gene (Fig. 5). Pairwise comparison of this region clearly shows the conservation of certain nucleotide sequences after considerable divergence. Both IFN-α and IFN-β can be produced even from a single cell line upon induction by viruses, although only IFN-β can be produced after induction by poly(I)·poly(C) (40). In view of the fact that both genes are derived from a common ancestor and that they are induced by the common (or related) inducers, it is tempting to speculate that this conserved region plays a role in the regulation of IFN gene expression. Moreover, within this conserved region, one can find a notable bias for pyrimidines over purines in the coding strand from nucleotide position -134 to -167. How this region is involved in gene expression with various inducers remains to be clarified.

We thank T. Maniatis, S. Kobayashi, and J. Beard for providing us with the human gene library, induced DIP 2 cells, and reverse transcriptase, respectively. We are indebted to H. Sugano, M. Muramatsu, and Y. Fujii-Kuriyama for support and suggestions. Thanks are also due to J. Vilček, R. Kominami, Y. Mishima, Y. Mizukami, Y. Urano, K. Noguchi, and C. Takaoka for help. Kai Zinn is also greatly acknowledged for the preliminary phase of this work. We are grateful to S. Ichikawa for typing the manuscript.

1. Taniguchi, T., Fujii-Kuriyama, Y. & Muramatsu, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4003–4006.
2. Knight, E., Jr., Hunkapiller, M. W., Korant, B. D., Hardy, R. W. F. & Hood, L. E. (1980) *Science* **207**, 525–526.
3. Sehgal, P. B. & Sagar, A. D. (1989) *Nature (London)* **288**, 95-97.
4. Weissenbach, J., Chernajovsky, Y., Zeevi, M., Shulman, L., Soreq, H., Nir, U., Wallach, D., Perricaudet, M., Tiollais, P. & Revel, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7152–7156.
5. Taniguchi, T., Sakai, M., Fujii-Kuriyama, Y., Muramatsu, M., Kobayashi, S. & Sudo, T. (1979) *Proc. Jpn. Acad.* **55B**, 464–469.
6. Taniguchi, T., Ohno. S., Fujii-Kuriyama, Y. & Muramatsu, M. (1980) *Gene* **10**, 11–15.
7. Houghton, M., Stewart, A. G., Deel, 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. & Carey, N. H. (1980) *Nucleic Acids Res.* **8**, 1913–1931.
8. Houghton, M., Eaton, M. A. W., Stewart, A. G., Smith J. C., Deel, S. M., Catlin, G. H., Lewis, H. M., Patel, T. P., Emtage, J. S., Carey, N. H. & Porter, A. G. (1980) *Nucleic Acids Res.* **8**, 2885–2894.
9. Derynck, R., Content, J., DeClercq, E., Volckaert, G., Tavernier, J., Devos, R. & Fiers, W. (1980) *Nature (London)* **285**, 542–547.
10. Nagata, S., Taira, A., Hall, L., Johnsrud, M., Streuli, J., Ecsödi, W., Boll, K., Cantell, C. & Weissmann, C. (1980) *Nature (London)* **284**, 316–320.
11. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. & Weissmann, C. (1980) *Gene* **10**, 1–10.
12. Streuli, M., Nagata, S. & Weissmann, C. (1980) *Science* **209**, 1343–1347.
13. Taniguchi, T., Guarente, L., Roberts, T. M., Kimelman, D., Douhan, J., III & Ptashne, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5230–5233.
14. Derynck, R., Remaut, E., Saman, E., Stanssens, P., DeClercq, E., Content, J. & Fiers, W. (1980) *Nature (London)* **287**, 193–197.
15. Goeddel, D. V., Shepard, H. M., Yelverton, E., Leung, D., Drea, R., Sloma, A. & Pestka, S. (1980) *Nucleic Acids Res.* **8**, 4057–4074.
16. Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Muramatsu, M. & Weissmann, C. (1980) *Nature (London)* **285**, 547–549.
17. Nagata, S., Mantei, N. & Weissmann, C. (1980) *Nature (London)* **287**, 401–408.
18. Gross-Bellard, M., Oudet, P. & Chambon, P. (1973) *Eur. J. Biochem.* **36**, 32–38.
19. Currier, T. C. & Nester, E. W. (1976) *Anal. Biochem.* **76**, 431–441.
20. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
21. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
22. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
23. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727–3731.
24. Roop, D. R., Nordstrom, J. L., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. (1978) *Cell* **15**, 671–685.
25. Kataoka, T., Yamawaki-Kataoka, Y., Yamagishi, H. & Honjo, T. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4240–4244.
26. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
27. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2398.
28. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
29. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12**, 721–732.
30. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **7**, 1175–1193.
31. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980) *Science* **209**, 1406–1413.
32. Hentshel, C., Irminger, J.-C., Bucher, P. & Birnstiel, M. L., (1980) *Nature (London)* **285**, 147–151.
33. Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Richards, J. E., Slightom, J. L., Tucker, P. W. & Smithies, O. (1978) *Science* **202**, 1279–1284.
34. Proudfoot, N., Shander, M. H. M., Manley, J. L., Gefter, M. D. & Maniatis, T. (1980) *Science* **209**, 1329–1335.
35. Hamer, D. H., Smith, K. O., Bayer, S. H. & Leder, P. (1979) *Cell* **17**, 725–735.
36. Lai, C.-J., Smith, K.O., Bayer, S. H. & Leder, P. (1979) *Cell* **17**, 725–735.
37. Aleström, P., Akusjärvi, G., Perricaudet, M., Matheus, M. B., Klessig, D. F. & Pettersson, U. (1980) *Cell* **19**, 671–681.
38. Schaffner, W., Kunz, G., Daetwyler, H., Telford, T., Smith, H. O. & Birnstiel, M. L. (1979) *Cell* **14**, 655–671.
39. Grosschedl, R. & Birnstiel, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7102–7106.
40. Pang, R. H. L., Hayes, T. G. & Vilček, J. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5341–5345.