Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5435–5439, September 1981
Biochemistry

DNA sequence of a major human leukocyte interferon gene

(gene sequence/introns/eukaryotic genes/*Alu* repeats)

RICHARD M. LAWN, MITCHELL GROSS, CATHERINE M. HOUCK, ARTHUR E. FRANKE, PATRICK V. GRAY, AND DAVID V. GOEDDEL

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080

Communicated by David Marshall Prescott, May 29, 1981

ABSTRACT **The gene for human leukocyte interferon α2 (designated either LeIF A or HuIFN-α2) has been isolated from a human genome library. The DNA sequence of this gene demonstrates that it lacks introns. The 3′ noncoding sequences of the IFN-α2 gene correspond to two types of IFN-α2 cDNA clones we have isolated that have alternate sites of polyadenylylation. A comparison of seven human IFN-α sequences shows that they are homologous in the 5′ flanking region and contain identical "TATA box" sequences. The recombinant λ clone containing the IFN-α2 gene also contains two copies of the "*Alu* family" repeat sequence.**

The human interferon genes compose a multigene family containing at least 10 leukocyte interferon (designated LeIF or IFN-α) genes (1–6), which share about 85–95% sequence homology (4), and a single fibroblast interferon (designated FIF or IFN-β1) gene, which shares about 50% DNA sequence homology with the IFN-α genes (7–9). These genes have been located on human chromosome 9 (10), and several of them have been shown to be very closely linked by their appearance on single λ Charon 4a human genome recombinant clones (3, 6). A second type of interferon cDNA clone (IFN-β2) has now been isolated from human fibroblasts (11). It bears little sequence homology to either IFN-β1 or the IFN-α genes, and its chromosome location has not been determined.

Many of the IFN-α genes, as well as hybrids between two different interferon genes, have now been expressed in *Escherichia coli* cells (1, 12–14). This has allowed testing of purified IFN-α proteins in various animal cells and against different virus challenges. These results indicate different activities for the various proteins and point to the potential usefulness of multiple interferon species. It is hoped that the study of this related multigene family at the amino acid and nucleotide sequence levels will help to identify features of importance for protein function and gene expression. In this report we present the gene sequence of the human IFN-α2 gene and compare it with its mRNA sequence established from cDNA clones and with other IFN-α gene sequences obtained by ourselves and by others.

MATERIALS AND METHODS

Isolation of Cloned DNA. A bacteriophage λ Charon 4a recombinant library constructed by Lawn *et al.* (15) was screened for clones containing IFN-α sequences with radioactive probe from the cDNA clone LeIF A(1) as described (6, 16). Preparation of DNA, inserting *Eco*RI fragments into the plasmid pBR325 (17), and subcloning were described previously (16). P1–EK1 containment as specified by National Institutes of Health guidelines was employed.

Restriction Enzyme Mapping. DNA was digested with restriction enzymes (New England BioLabs) under conditions given by the supplier. Electrophoresis, transfer to nitrocellulose filter paper, and hybridization procedures have been described (4). The *Alu* family probe BLUR 8 (18) was obtained from Carl Schmid. Mapping of restriction sites was accomplished by single and double restriction enzyme digests (6) and by the partial digestions procedure of Smith and Birnstiel (19).

DNA Sequence Analysis. Plasmid DNA was cleaved with restriction endonucleases and end-labeled with [γ-^{32}P]ATP (Amersham) by using polynucleotide kinase (P-L Biochemicals), and the sequence was determined by the methods of Maxam and Gilbert (20) with the exception that DNA fragments were purified from gel slices by electroelution (16).

RESULTS

Characterization of λα2. The phage λ Charon 4a/human genome recombinant designated λα2 was among those isolated from the human genome library of Lawn *et al.* (15), utilizing as radioactive probe the cDNA clone LeIF A (4, 6). Restriction endonuclease digestion revealed that λα2 contains a 16.3-kilobase pair (kb) insert of human DNA (Fig. 1). Blot hybridization (21) with the LeIF A probe showed that the interferon gene in λα2 is contained within a single *Eco*RI fragment of 4.3 kb, a HindIII fragment of >20 kb, and three *Bgl* II fragments of >20, 7.3, and 0.3 kb (Fig. 1.)

A restriction endonuclease map of λα2 is shown in Fig. 2. The map was constructed by digestion of the phage DNA with endonucleases singly or in combination, by isolation of individual digestion fragments with subsequent digestion by a second endonuclease, and by partial digestion of isolated fragments that were end labeled at the single *Hin*dIII site within the human DNA. Electrophoresis in both agarose and polyacrylamide gels allowed detection of DNA fragments of a wide size range. The direction of transcription of the interferon gene within this clone was determined by a combination of restriction mapping and direct DNA sequence analysis. The horizontal arrow above the gene in Fig. 2 indicates this orientation.

DNA Sequence Analysis. DNA fragments from a subclone of the 4.3-kb *Eco*RI fragment of λα2 were end labeled with ^{32}P at restriction endonuclease sites indicated in Fig. 2 and their sequences were determined by the method of Maxam and Gilbert (20). The DNA sequence of the IFN-α gene contained in λα2 and several hundred nucleotides flanking each side of the protein coding region is shown in Fig. 3. The sequence of the message coding region of the λα2 gene was compared to the sequence of eight distinct IFN-α cDNA clones we have isolated from the myeloid cell line KG-1 (4), and two cDNA clones obtained by others (2, 22). Within the translated region, the λα2 sequence is identical to that derived from the cDNA clone designated IFN-α2 (2) and differs by a single nucleotide in the cod-

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.

Abbreviations: bp, base pair(s); kb, kilobase pair(s); LeIF, leukocyte interferon; IFN, interferon.

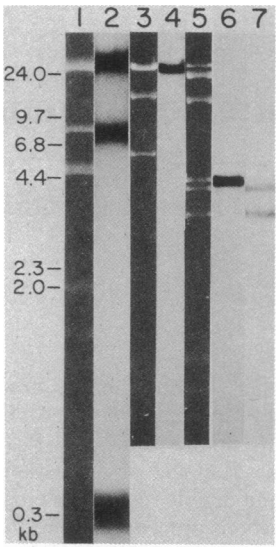

FIG. 1. Restriction digests and blot hybridizations. λα2 DNA (2 μg) was digested with restriction endonucleases, electrophoresed in 0.75% agarose, and stained with ethidium bromide. Lane 1, *Bgl* II; lane 3, *Hin*dIII; lane 5, *Eco*RI. After transfer to nitrocellulose the DNA was hybridized with ^{32}P-labeled DNA from the leukocyte interferon cDNA clone pLeIF A: lane 2, *Bgl* II; lane 4, *Hin*dIII; lane 6, *Eco*RI: In lane 7 *Eco*RI-digested DNA was hybridized with ^{32}P-labeled *Alu* repeat family DNA from the clone BLUR 8. An approximate size scale is indicated on the left. Restriction endonuclease fragments of λ and pBR322 served as standards in the gels. The sizes of the DNA fragments that hybridize to the interferon cDNA probe are: *Bgl* II, >20, 7.3, and 0.3 kb; *Hin*dIII, >20 kb; *Eco*RI, 4.3 kb. The *Eco*RI fragments that hybridize to the BLUR 8 probe are 4.0 and 3.3 kb.

ing region from LeIF A (1, 4). The single coding nucleotide difference between LeIF A and both λα2 and IFN-α2 creates an amino acid substitution of arginine for lysine at the amino acid number 23 of the mature protein. This may reflect the nature of this gene in the KG-1 cell line.

In the untranslated regions, the only nucleotide differences between λα2 and the LeIF A cDNA are two single-base changes in the 3′ noncoding region that extends to the site of poly(A) addition in LeIF A (indicated by an arrow at nucleotide 900 in Fig. 3) and a small region of the 5′ terminus of LeIF A located between the asterisks at positions −61 and −49 in Fig. 3. The discrepancy in this 5′ region may reflect improper base pairing that occurred during the initiation of second-strand DNA synthesis in the construction of the cDNA clone LeIF A. [Neither the cDNA clone IFNα2 (2) nor any other copies of LeIF A that we obtained (4) extended this far in the 5′ direction.] The presumed true 5′ terminus of this particular mRNA species is located within two nucleotides of position −68 (indicated by arrow in Fig. 3) by analogy with other IFN-α mRNAs (3, 6). The colinearity between the gene and cDNA sequence demonstrates that the λα2 gene is devoid of introns.

The sequence T-A-T-T-T-A-A is found 32 nucleotides before the presumed mRNA cap site of the λα2 gene. This sequence may correspond to the Goldberg–Hogness, or TATA, box sequence located about 30–40 nucleotides 5′ to the cap site of many eukaryotic genes transcribed by RNA polymerase II (23–25). A comparison of the 5′ noncoding regions of seven IFN-α genes is shown in Fig. 4 and is discussed below.

The sequence of the 3′ noncoding region of the λα2 gene contains the A-A-T-A-A-A hexanucleotide beginning 27 nucleotides before the site of poly(A) addition of the cDNA clone LeIF A (1). This sequence, or a close variant, precedes the polyadenylylation site of most eukaryotic genes (26). In our characterization of IFN-α cDNA clones isolated from KG-1 cells, a clone was found that was identical in DNA sequence to LeIF A (including the A-A-T-A-A-A hexanucleotide) but continued for an additional 175 nucleotides before termination with poly (A). A second A-A-T-A-A-A hexanucleotide precedes the 3′ end of this extended cDNA clone [designated LeIF A(+175)] by 20

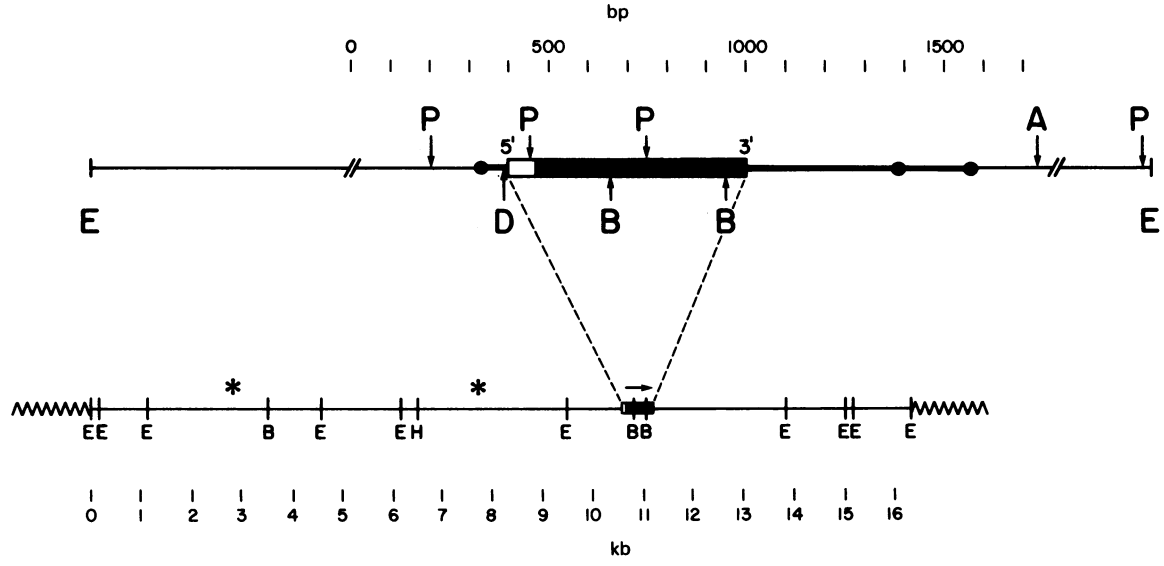

FIG. 2. Restriction endonuclease maps. The sites of cleavage with restriction endonucleases *Eco*RI (E), *Bgl* II (B) and *Hin*dIII (H) are shown in the lower part of the figure for the human DNA portion of the recombinant phage λα2. Wavy lines indicate the λ Charon 4A arms. The location of the interferon gene is indicated by a filled box. The arrow above the gene denotes the direction of transcription. Asterisks indicate restriction fragments containing *Alu* repeat family sequences. A size scale in kb is shown below. An enlargement of the gene region is shown above with a size scale in base pairs (bp). DNA sequence determination proceeded from indicated *Pvu* II (P), *Bgl* II (B), *Acc* I (A), and *Dde* I (D) sites. Not all *Dde* I and *Acc* I sites in this region are shown. The coding sequence for the mature interferon protein is indicated by the filled box and the signal peptide by the open box. Beyond the 3′ end of the gene, the noncoding region is indicated by a heavy line terminating in the poly(A) sites of cDNA clones LeIF A and LeIF A (+175) which are shown by filled circles. The 5′ noncoding region and presumed cap site of the mRNA are also indicated by a heavy line and filled circle.

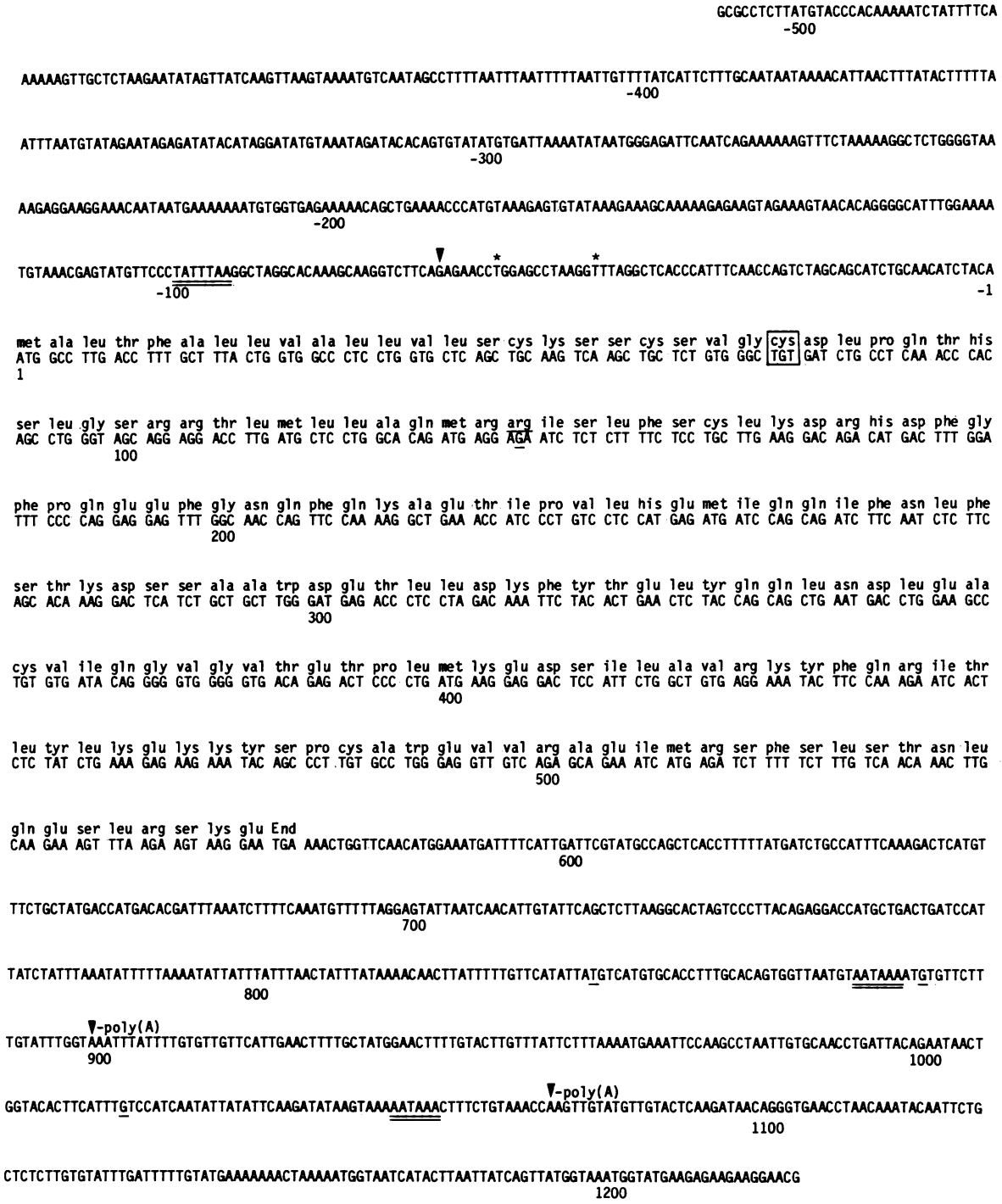

FIG. 3. DNA sequence of the $\lambda\alpha 2$ interferon gene and flanking regions. The DNA sequence of the interferon gene containing portion of the recombinant phage $\lambda\alpha 2$ was determined by the method of Maxam and Gilbert (20). Nucleotides have been numbered from +1 at the first translated nucleotide and −1 from the preceding nucleotide. The predicted amino acid sequence is shown above the nucleotide sequence. Nucleotides and the single amino acid that differ from the cDNA clone LeIF A (1) are underlined. Asterisks denote a short nonhomologous region at the 5′ end of the LeIF A clone. The initial amino acid of the mature protein is enclosed in a box. The 5′ TATA (Goldberg–Hogness) box and 3′ polyadenylylation signal sequences are underlined twice. Arrows mark the presumed 5′ (±2 nucleotides) and 3′ termini of corresponding mRNAs. See text for further details.

nucleotides. The additional 175 nucleotides at the 3′ end of LeIF A(+175) are the same as the 3′ flanking nucleotides of the $\lambda\alpha 2$ gene, with the exception of one base change. Hence, the LeIF A(+175) cDNA clone may have arisen from an mRNA species that continued transcription beyond one possible polyadenylylation site to a second site further downstream.

Clustering of LeIF Genes. All of the 10 or more human IFN-α genes and the IFN-β1 (fibroblast interferon) gene detectable by cross-hybridization have been assigned to human chromosome 9 by blot hybridizations of cDNA probes to DNA obtained from a series of mouse–human hybrid cell lines (10). At least some of these genes are very closely linked. We have characterized two λ Charon 4a recombinant clones containing pairs of IFN-α genes separated by 12.3 kb and 5.0 kb (ref. 6; un-

published results). Nagata *et al.* (3) have also reported two λ recombinant clones isolated from the same human genome library that contain pairs of IFN-α genes. To determine whether the λα2 gene is closely linked to other IFN-α genes, we tested this λ clone for overlap with other IFN-α-containing λ clones that we have thus far isolated. Terminal *Eco*RI fragments of 0.2 and 0.22 kb were isolated from λα2 DNA. These fragments lie at opposite ends of the human DNA insert of this recombinant (Fig. 2). These fragments were ^{32}P labeled and used as hybridization probes to Southern blots of six λIFN-α clones. The probes hybridized only to the λα2 DNA from which they were derived, indicating no overlap with the other five clones.

***Alu* Family Repeat Sequences.** The human genome contains more than 300,000 copies of a family of related sequences approximately 300 bp in length termed the *Alu* family (27). Although this sequence shares a short region of homology with papovavirus origins of replication and small nuclear RNAs (28), no function for the *Alu* repeats has been demonstrated. Copies of the *Alu* family sequence have been found in the vicinity of the β-globin-like genes (28, 29) and the insulin gene (30). Using an *Alu* family hybridization probe (BLUR 8; ref. 18), we detected two copies of this repeat upstream of the λα2 gene. The *Eco*RI fragments hybridizing to BLUR 8 can be seen in lane 7 of Fig. 1. The location of these hybridizing fragments is indicated by asterisks in Fig. 2. DNA sequence analysis of these *Alu* family sequences and those in other λIFN-α clones will be reported in a manuscript to follow.

DISCUSSION

The human interferon genes are members of a related multigene family. The study of gene families should contribute to the understanding of gene control and of functional relationships among related genes and proteins, as well as the processes of evolution by gene duplication and divergence. Interferons of slightly different protein sequence may respond differently to various inducers, modulate various responses, be expressed preferentially in certain cells, or recognize different target cell types. The latter possibility is consistent with the findings that individual human interferon proteins produced by recombinant DNA methods display varied activity in different animals and on cultured cells derived from various organs (31). Different cloned interferon proteins, including hybrid proteins, may also find specific clinical applications.

The members of the IFN-α and IFN-β1 family may be genetically linked. These genes have been assigned to human chromosome 9 (10). Several λ recombinant clones with pairs of linked IFN-α genes have been reported (3, 6). Because λα2 does not overlap existing recombinant clones, it remains to be seen whether all the members of the human interferon gene family are clustered.

The interferon gene reported here is devoid of introns, in contrast to most eukaryotic nuclear genes whose sequences have been determined to date. However, the absence of introns has been observed in all of the six other IFN-α (3, 6) as well as the IFN-β1 (16) interferon genes that have been analyzed in detail.

All of the approximately 12 distinct human IFN-α genes whose sequences have been determined to date share approximately 85–95% nucleotide sequence homology in the protein-coding region. Predicted amino acid variations occur throughout the gene without extreme clustering, although several areas of possible conservation occur (4). The nucleotide sequence homology between flanking regions of IFN-α genes is less than in the coding regions, but is still substantial. Such homologies may reflect sequences of functional importance, but they may also result from recent episodes of gene duplication or conversion, as were postulated for paired globin genes (32, 33). Fig. 4 presents the sequence alignment of approximately 200 nucleotides preceding the translation initiator for the six IFN-α genes we have analyzed to date (6) and one gene (IFN-α1) reported by Nagata *et al.* (3). The overall sequence homology in this region is about 75%. The presumed 5′ terminus of the IFN-α mRNAs is within several nucleotides of the position marked by an arrow in the figure (3, 6). Preceding this cap site by about 30 nucleotides is the sequence 5′-T-A-T-T-T-A-A, which is conserved exactly in all six of these IFN-α genes. This presumably represents the TATA box sequence that precedes many eukaryotic genes transcribed by RNA polymerase II (23–25). Alteration of the TATA sequence in other genes has been shown to modulate the level and specificity of transcription *in vitro* and *in vivo* (25, 34–38). The CAAT motif existing about 40 nucleotides further upstream in histone (37) and in globin (24) genes and found to modulate levels of transcription is absent from the interferon genes. Any further speculation on the role of 5′ interferon sequences in the control of gene expression must await direct experiments involving *in vitro* mutagenesis and transcription in surrogate systems.

```
1  ATGAAAAAAATGGTGAGAAAAACAGCTGAAAACCCATGTAAAGAGTGTAT-AAAGAAAGCAAAAAGAGAAGTAG-AAAGTAACACAGGGGCATTTGGAAAATGTAAACGAGTA
2  AAATAAAAGTGGTTGAG--AAACTGCTCTACACCCATGTAGACAGGACATAAAGGAAAGCCAAAAGAGAAGTAGAAAAAAACATGAAGAAGCTTCAGAAAATGGAAGCTAGTA
3  TGATGAAAAACAATGAA--AAACATTCTTAAACACATGTAGAGAGTGCAA-AAAGAAAGCAAAAACAGACATAG-AAAGTAAAACTAGGGCATTTAGAAAATGGAAATTAGTA
4  ATGAAAAAAAAAATGAA--AAACGTATTTAAACACATGGAGAGAGTGCAT-AAAGAAAGCAAAAACAGAGATAG-AAAGTAAAACTAGGGCATTTAGAAAATGGAAATTAGTA
5  ATGATGAAAAAAAAACGA--AAAACTTTTTAAACACATGGAGAGAGTACAT-AAAGAAAGCAAAAACAGAGATAG-AAAGTAAAACTAGGGCATTTAGAAAATGGAAATTAGTA
6  AAAACAAAACATTTGAGAAACACGGCTCTAAACTCATGTAAAGAGTGCATGAAGGAAAGCAAAAACAGAAATGG-AAAGTGGCCCAGAAGCATTAAGAAAGTGGAAATCAGTA
7  ACACACAAAACTGGTTGAAAAACTACTCTATACCCATGTAGAGAGTAAATAAATGAAAGCAAAATCAGACGTAG-AAAGTAAA---------TTCTGAAAATGGAAACTAGTA
```

```
1  TGTTCCCTATTTAAGGC-TAGGCACAAAGCAAGGTCTTCAGAGAACCTGGAGCCTAAGGTTTAGGCTCACCCATT-TCAACCAGTCTAGCAGCATCTGCAACATCTACAATG
2  TGTTCCTTATTTAAGACCTATGCACAGAGCAAGGTCTTCAGAAAACCTACAACCCAAGGTTCAGTGTTACCCCTCATCAACCAGCCCAGCAGCATCTTCAGGGTTCCCAATG
3  TGTTCACTATTTAAGGCCTATGCACAGAGCAAAGTCTTCAGAAAACCTAGAGGCCAAAGTTCAAGGTTACCCATC-TCAAGTAGCCTAGCAACATTTGCAACATCCCA-ATG
4  TGTTCACTATTTAAGACCTATGCACAGAGCAAAGTCTCCAGAAAACCTAGAGGCCACGGTTCAA-GTTACCCACC-TCAGGTAGCCTAGTGATATTTGCAAAATCCCA-ATG
5  TGTTCACTATTTAAGACCTATGCACAGAGCAAAGTCTTCAGAAAACCTAGAGGCCGAAGTTCAAGGTTATCCATC-TCAAGTAGCCTAGCAATATTTGCAACATCCCA-ATG
6  TGTTCCCTATTTAAGGCATTTGCAGGAAGCAAGGCCTTCACAGAACCTAGAGCCCAAGGTTCAGAGTCACCCATC-TCAGCAAGCCCAGAAGTATCTGCAATATCTACGATG
7  TGTTCCCTATTTAAGACCTACACATAAAGCAAGGTCTTCAGAGAACCTAGAGCTGAAGGTTCAGAGTCACCCATC-TCAACAAGTCCAACAGCATCTGCAACATCTACAATG
```

FIG. 4. Comparison of 5′ sequences. The initial 120 nucleotides preceding the start of translation (*ATG*) of seven human leukocyte interferon genes (IFN-α) are compared. Several gaps (−) were introduced to maximize alignment. The TATA box is enclosed. An arrowhead points to the presumed cap site (±2 nucleotides) of corresponding mRNA species. Names and references of the gene sequences are: 1, λ4a (this paper); 2, λ2h (6); 3, $\lambda 2c_1$ (6); 4, λ1j (unpublished data); 5, λ1l (unpublished data); 6, IFN-α1 (3); 7, λ5k (unpublished data).

We thank John Adelman for excellent technical assistance and Tom Maniatis and Carl Schmid for providing the human λ recombinant library and BLUR 8 cDNA clone, respectively. This research was supported by Genentech, Inc., and Hoffmann–La Roche, Inc.

1. 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. & Pestka, S. (1980) *Nature (London)* **287**, 411–416.
2. Streuli, M., Nagata, S. & Weissmann, C. (1980) *Science* **209**, 1343–1347.
3. Nagata, S., Mantei, N. & Weissmann, C. (1980) *Nature (London)* **287**, 401–408.
4. Goeddel, D. V., Leung, D. W., Dull, T. J., Gross, M., Lawn, R. M., McCandliss, R., Seeburg, P. H., Ullrich, A., Yelverton, E. & Gray, P. W. (1981) *Nature (London)* **290**, 20–26.
5. Allen, G. & Fantes, K. H. (1980) *Nature (London)* **287**, 408–413.
6. Lawn, R. M., Adelman, J., Dull, T. J., Gross, M., Goeddel, D. V. & Ullrich, A. (1981) *Science* **212**, 1159–1162.
7. Taniguchi, T., Ohno, S., Fujii-Kurimaya, Y. & Muramatsu, M. (1980) *Gene* **10**, 11–15.
8. Derynk, R., Content, J., De Clerq, E., Volckaert, G., Tavernier, J., Devos, R. & Fiers, W. (1980) *Nature (London)* **285**, 542–547.
9. Goeddel, D. V., Shepard, H. M., Yelverton, E., Leung, D. & Crea, R. (1980) *Nucleic Acids Res.* **8**, 4057–4074.
10. Owerbach, D., Rutter, W. J., Shows, T. B., Gray, P., Goeddel, D. V. & Lawn, R. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3123–3127.
11. Weissenbach, J., Chernajovsky, Y., Zeevi, M., Shulman, L., Soreq, H., Nir, V., Wallach, D., Perricaudet, M., Tiollais, P. & Revel, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7152–7156.
12. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Escodi, J., Boll, W., Cantell, K. & Weissmann, C. (1980) *Nature (London)* **284**, 316–320.
13. Yelverton, E., Leung, D., Weck, P., Gray, P. W. & Goeddel, D. V. (1981) *Nucleic Acids Res.* **9**, 731–741.
14. Streuli, M., Hall, A., Boll, W., Stewart, W. E., II, Nagata, S. & Weissmann, C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2848–2852.
15. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
16. Lawn, R. M., Adelman, J., Franke, A. E., Houck, C. M., Gross, M., Najarian, R. & Goeddel, D. V. (1981) *Nucleic Acids Res.* **9**, 1045–1052.
17. Bolivar, F. (1978) *Gene* **4**, 121–134.
18. Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T. & Schmid, C. W. (1980) *Nature (London)* **284**, 372–374.
19. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2397.
20. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
21. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
22. Mantei, N., Schwartzsteen, M., Streuli, M., Panem, S., Nagata, S. & Weissmann, C. (1980) *Gene* **10**, 1–10.
23. Goldberg, M. (1979) Dissertation (Stanford Univ., Stanford, CA).
24. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.
25. Grosschedl, R. & Birnstiel, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1432–1436.
26. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
27. Houck, C. M., Rinehart, F. P. & Schmid, C. W. (1979) *J. Mol. Biol.* **132**, 289–306.
28. Jelinek, W. R., Toomey, T. P., Leinwand, L., Duncan, C. H., Biro, P. A., Choudary, P. V., Weissman, S. M., Rubin, C. M., Houck, C. M., Deininger, P. L. & Schmid, C. W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1398–1402.
29. Fritsch, E. F., Lawn, R. M. & Maniatis, T. (1980) *Cell* **19**, 959–972.
30. Bell, G. I., Pictet, R. & Rutter, W. J. (1980) *Nucleic Acids Res.* **8**, 4091–4109.
31. Stebbing, N., Weck, P. K., Fenno, J. T., Apperson, S. & Lee, S. H. (1981) in *The Biology of the Interferon System*, eds. DeMaeyer, E., Galasso, G. & Schellekens, H. (Elsevier, Amsterdam), in press.
32. Lauer, J., Shen, C. J. & Maniatis, T. (1980) *Cell* **20**, 119–130.
33. Slightom, J. L., Blechl, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
34. Benoist, C. & Chambon, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3865–3869.
35. Wasylyk, B., Kedinger, C., Corden, J., Brison, O. & Chambon, P. (1980) *Nature (London)* **285**, 367–373.
36. Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Teoule, R. & Chambon, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7024–7028.
37. Grosschedl, R. & Birnstiel, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7102–7106.
38. Mathis, D. I. & Chambon, P. (1981) *Nature (London)* **290**, 310–315.