The EMBO Journal vol.5 no.7 pp.1601–1606, 1986

Characterization of a human gene inducible by α- and β-interferons and its expression in mouse cells

John M.Kelly[1], Andrew C.G.Porter, Yuti Chernajovsky, Christopher S.Gilbert, George R.Stark and Ian M.Kerr

Imperial Cancer Research Fund Laboratories, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

[1]Present address: Wolfson Unit of Molecular Medical Microbiology and Parasitology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Communicated by G.R.Stark

An intact interferon-inducible gene has been isolated from a cosmid library of human genomic DNA. The gene (designated 6-16) encodes a mRNA of approximately 1 kb which is induced well by α- and β- but poorly by γ-interferons. Genomic and cDNA sequences indicate that the gene contains five exons, and that the mRNA encodes a hydrophobic polypeptide of 130 amino acids with a putative NH_2-terminal signal sequence. The 5′ end has been identified by primer extension. The corresponding genomic DNA contains a TATA box 20 nucleotides upstream of the putative transcription initiation site. After transfection of the human genomic cosmid into mouse Ltk⁻ cells, human 6-16 mRNA is expressed in response to mouse α- and β- but not γ-interferons with the same kinetics and dose-response as in the human cells. No such expression is observed in response to human interferons. It can be concluded that the human cosmid DNA contains all of the sequences necessary for α- and β-interferon-induced gene expression and that the mechanisms governing such expression are conserved between murine and human cells.

Key words: gene/human/interferon-inducible/transfection/expression

Introduction

Several groups have isolated and characterized complementary DNA (cDNA) clones corresponding to mRNAs induced by interferons in human (Friedman *et al.*, 1984; Larner *et al.*, 1984; Benech *et al.*, 1985; Luster *et al.*, 1985) and mouse (Samanta *et al.*, 1984; Staeheli *et al.*, 1986) cells. Induction by the α-interferons at least is mediated both transcriptionally and post-transcriptionally (Friedman *et al.*, 1984; Larner *et al.*, 1984) and it appears that α- and γ-interferons can differentially regulate the levels of some mRNAs (Rosa *et al.*, 1983b; Kelly *et al.*, 1985; Strunk *et al.*, 1985). Homologous sequences have been noted in the 5′-flanking regions of a number of interferon-inducible HLA Class I and II genes and the interferon-inducible gene for human metallothionein II (Friedman and Stark, 1985). Neither the role of these sequences nor the mechanisms by which the different interferons (α, β and γ) regulate gene expression are understood. For this a more detailed analysis of a variety of different interferon-inducible genes and the factors controlling their expression will be required.

The gene 6-16 encodes an mRNA of approximately 1 kb that is highly induced by α-interferons in a variety of human cells (Friedman *et al.*, 1984; Kelly *et al.*, 1985). After induction 6-16 mRNA can constitute as much as 0.1% of the total mRNA. Induction is a primary response independent of protein synthesis (Kelly *et al.*, 1985). Using an antibody to the interferon-inducible membrane protein P16 (Burrone and Milstein, 1982) Evans, Secher and Milstein (personal communication) have isolated a cDNA, the sequence of which is essentially identical to that for 6-16 reported here.

The 6-16 gene is poorly induced, if at all, in response to γ-interferon (Kelly *et al.*, 1985), as is the case for the mouse Mx protein gene (Staeheli *et al.*, 1986). In contrast there are genes inducible predominantly by γ-interferon and those which respond well to α-, β- or γ-interferons (Luster *et al.*, 1985; Strunk *et al.*, 1985; Benech *et al.*, 1985). A comparison of the results obtained on detailed analysis of these differently inducible genes should provide a basis for determining both the differences in the mechanisms by which Type I (α,β) and Type II (γ) interferon receptors communicate with the nucleus and the DNA sequences involved in the control of α-, β- and γ-inducible gene expression.

Results

Induction of 6-16 mRNA by Type I interferons

6-16 mRNA is highly induced by α-interferons in all human cell

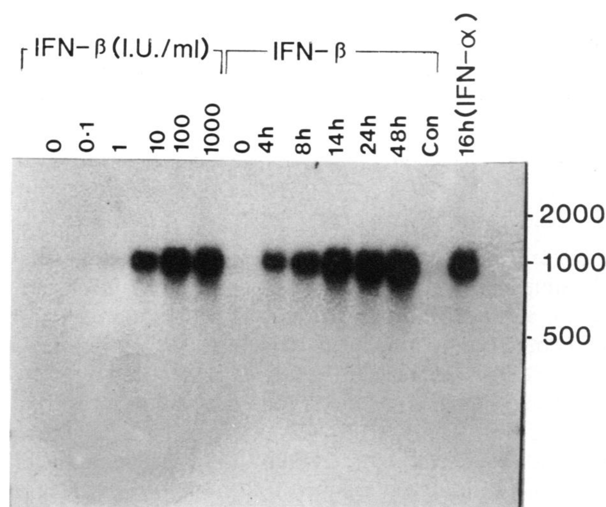

Fig. 1. Induction of 6-16 mRNA by β-interferon. A Northern analysis (see Materials and methods) of the dose-response **(tracks 1–6)** and kinetics **(tracks 7–12)** of induction of 6-16 mRNA by β-interferon in subconfluent monolayer cultures of HeLa cells is presented. β-Interferon treatment was either for 16 h at 37°C with the indicated dose **(tracks 1–6)** or with 300 IU/ml for the indicated time **(tracks 7–12)**. **Tracks 13** and **14:** RNA from cells incubated for 16 h at 37°C in the absence **(track 13)** or presence **(track 14)** of α-interferons (300 IU/ml). Aliquots (20 μg) of total cellular RNA were analysed (see Materials and methods). The positions to which size markers (in nt) migrated are indicated to the right.

© IRL Press Limited, Oxford, England

```
GENOMIC
GAACCGTTTACTCGCTGCTGTGCCCATCTATCAGCAG

CDNA                                                                                                                   120
GCTCCGGGCTGAAGATTGCTTCTCTTCTCTCCTCCAAGGTCTAGTGACGGAGCCCGCGCGCGCCCACCATGCGGCAGAAGGCGGTATCCGTTTTCTTGTGCTACCTGCTGCTCTTCACT
                                                                     MetArgGlnLysAlaValSerValPheLeuCysTyrLeuLeuLeuPheThr
                                                                                                                       240
TGCAGTGGGGTGGAGGCAGGTAAGAAAAAGTGCTCGGAGAGCTCGGACAGCGGCTCCGGGTTCTGGAAGGCCCTGACCTTCATGGCCGTCGGAGGAGGACTCGCAGTCGCCGGGCTGCCC
CysSerGlyValGluAlaGlyLysLysLysCysSerGluSerSerAspSerGlySerGlyPheTrpLysAlaLeuThrPheMetAlaValGlyGlyGlyLeuAlaValAlaGlyLeuPro
                                                                                                                       360
GCGCTGGGCTTCACCGGCGCCGGCATCGCGGCCAACTCGGTGGCTGCCTCGCTGATGAGCTGGTCTGCGATCCTGAATGGGGGCGGCGTGCCCGCCGGGGGGCTAGTGGCCACGCTGCAG
AlaLeuGlyPheThrGlyAlaGlyIleAlaAlaAsnSerValAlaAlaSerLeuMetSerTrpSerAlaIleLeuAsnGlyGlyGlyValProAlaGlyGlyLeuValAlaThrLeuGln
                                                                                                                       480
AGCCTCGGGGCTGGTGGCAGCAGCGTCGTCATAGGTAATATTGGTGCCCTGATGCGGTACGCCACCCACAAGTATCTCGATAGTGAGGAGGATGAGGAGTAGCCAGCAGCTCCCAGAACC
SerLeuGlyAlaGlyGlySerSerValValIleGlyAsnIleGlyAlaLeuMetArgTyrAlaThrHisLysTyrLeuAspSerGluGluAspGluGlu-
                                                                                                                       600
TCTTCTTCCTTCTTGGCCTAACTCTTCCAGTTAGGATCTAGAACTTTGCCTTTTTTTTTTTTTTTTTTTTTTTGAGATGGGTTCTCACTATATTGTCCAGGCTAGAGTGCAGTGGCTATT
                                                                                                                       720
CACAGATGCGAACATAGTACACTGCAGCCTCCAACTCCTAGCCTCAAGTGATCCTCCTGTCTCAACCTCCCAAGTAGGATTACAAGCATGCGCCGACGATGCCCAGAATCCAGAACTTTG

TCTATCACTCTCCCCAACAACCTAGATGTGAAAACAGAATAAACTTCACCCAGAAAA
```

Fig. 2. Genomic and cDNA sequences corresponding to 6-16 mRNA and inferred peptide sequence. The entire sequence proposed as corresponding to intact 6-16 mRNA is represented by the 777 nt of cDNA sequence (see Materials and methods) plus 37 additional 5′ nt derived from genomic sequence based on primer extension and related experiments (Figure 5). The initiation codon of the longest open reading frame (cDNA residues 70–459) is preceded by CCACC (boxed), a eukaryotic initiation site consensus sequence (Kozak, 1984). The polyadenylation signal AATAAA (cDNA residues 756–761) is arrowed. Direct repeat sequences (cDNA residues 513–531 and 705–723) flanking the Alu repeat element are marked by broken underlines. The positions at which intron sequences occur in the corresponding genomic DNA are marked (▼, at cDNA residues 38, 139, 217 and 367). The inferred polypeptide sequence is shown and the putative NH_2-terminal signal sequence is underlined. If this sequence is indeed cleaved the most likely sites would be between Gly (20) and Val (21) or Ala (23) and Gly (24): each is about equally favoured by comparison with consensus data (von Heijne, 1984, 1985).

lines tested to date, including T98G neuroblastoma (Friedman *et al.*, 1984), HeLa (Kelly *et al.*, 1985), Daudi lymphoblastoid (McMahon *et al.*, 1986), U937 macrophage and F2001 and MRC5 fibroblasts (J.M.Kelly, data not shown). In HeLa cells at least it is also induced by β-interferon with the same kinetics and dose-response (Figure 1) as previously reported for the α-interferons (Kelly *et al.*, 1985). The appearance of this mRNA in significant amounts seems to be confined to cells treated with Type I interferons. It is present at only very low levels in the absence of interferon and, unlike the HLA Class I or 2-5A synthetase mRNAs, its level in HeLa cells is not affected by confluency. Nor, in contrast to metallothionein II, is it induced by heat shock (Kelly *et al.*, 1985; J.M.Kelly, unpublished data).

The sequence and characteristics of 6-16 mRNA

The complete sequence [814 nucleotides (nt) excluding the poly(A) tail] of 6-16 mRNA is represented in Figure 2 by 777 nt of sequence from 6-16 cDNA clones plus 37 nt of genomic sequence at the 5′ end (see below). The initiation codon of the longest open reading frame in 6-16 mRNA (390 nt, residues 70–459 in the cDNA sequence in Figure 2) is preceded by CCACC, a consensus sequence for eukaryotic translation initiation sites (Kozak, 1984). The first 20–23 amino acids of the inferred amino acid sequence (Figure 2) constitute a region characteristic of a signal peptide (von Heijne, 1984, 1985). Examination of the amino acid sequence and hydrophobicity plots (Kyte and Doolittle, 1982; data not presented) indicates that 6-16 is generally hydrophobic [there is no charged amino acid between residues 41 (Ala) and 115 (Met)], consistent with it being a membrane protein with a highly charged COOH-terminal tail.

An unusual, although not unique, feature of 6-16 mRNA is the presence in the 320 nt 3′-untranslated region of a sequence of approximately 175 nt sharing homology with Alu repeat elements (reviewed by Sharp, 1983). This sequence contains 23 consecutive thymidine residues at its 5′ end and is flanked by a direct repeat (Figure 2, 16 of the 19 nt, dashed lines) characteristic of integration sites in genomic DNA. A computer search of the NIH and EMBL databases revealed no other significant homologies between 6-16 mRNA and published sequences.

EcoRI BamHI HindIII
4·0kb → 1·8kb → ← 5·3kb ← 4·4kb

Fig. 3. Southern analysis of human genomic DNA with a 6-16 cDNA probe. Human genomic DNA (20 μg) digested with the indicated restriction enzymes was fractionated by electrophoresis through a 0.8% agarose gel and hybridized with a single-stranded M13 probe containing a 700 nt 6-16 cDNA insert (see Materials and methods). The size and position of genomic DNA fragments hybridizing to the probe are indicated. Human placental DNA **(track 1)** and Daudi cells **(tracks 2** and **3)** were analysed. Essentially identical results were obtained on similar analyses of HeLa cell DNA and that from the human placental DNA cosmid clones.

Structure of the 6-16 gene

Simple patterns indicative of a single gene were observed when Southern transfers of restricted human genomic DNA were probed with a 700 nt 6-16 cDNA (Figure 3). Consistent with this, clones corresponding to only one gene were isolated from a human placental DNA cosmid library (Hohn and Collins, 1980; Lindenmaier *et al.*, 1982). Approximately 300 000 colonies were

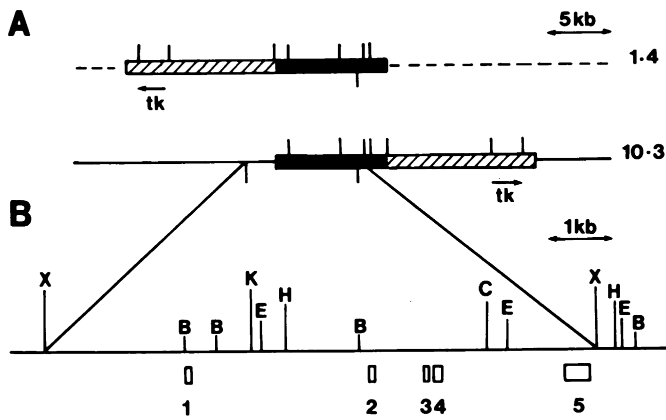

Fig. 4. Restriction maps of cosmid DNAs. **(A)** Overlapping regions (black boxes) of human insert DNA in cosmids 1.4 and 10.3 are shown. The hatched boxes represent vector DNA containing the HSV tk gene. The map of 10.3 for *Eco*RI (indicated above the line) and *Xho*I (below the line) sites is complete, the solid line representing the remainder of the human genomic DNA insert. The map for 1.4 is not complete and the remaining unmapped insert is represented by the broken line. **(B)** Enlargment of the region of cosmid 10.3 carrying the 6-16 gene showing the position of the five exons (open boxes) numbered from the 5′ end of the gene. Restriction enzyme sites for *Bam*HI (B), *Cla*I (C), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K) and *Xho*I (X) are indicated.

screened (see Materials and methods) and three distinct clones hybridizing to 6-16 cDNA were isolated. Southern analysis of the corresponding cosmid DNAs revealed that they contained overlapping inserts of 25–30 kb, and that in each case the 6-16 cDNA probe hybridized to the same restriction fragments as found with genomic DNA. One of the cosmids (coded 10.3) was shown, by transfection into mouse cells, to contain an intact interferon-inducible gene (see below); another (coded 1.4) was inactive in the transfection assay. These cosmids were analysed in more detail.

Limited restriction maps of both cosmids are presented (Figure 4A). Sequence analysis of appropriate genomic fragments of 10.3 and comparison with the known cDNA sequence showed that the gene contains five exons (Figure 4B). In addition cosmid 10.3 contains at least 20 kb of human DNA 5′ to the 6-16 gene (Figure 4A). The overlapping cosmid 1.4 does not; it terminates in the 0.5 kb *Bam*HI fragment which encodes the first exon, while extending substantially into the 3′ flanking region (Figure 4A). The genomic exon sequences in 10.3 were identical to those obtained by sequencing the cDNAs with one exception. In the cDNA clones, all of which were derived from Daudi cell mRNA, the oligo(dT) sequence in the Alu repeat contained 23 nt, whereas it was 25 nt long in the genomic clones of placental origin. The sequences at each intron–exon boundary (Table IA) and that surrounding the polyadenylation site (Table IB) are in good agreement with consensus sequences for such sites (see Discussion).

The transcription start site and adjacent 5′ flanking sequence

The 5′ end of the 6-16 mRNA was identified by primer extension. A synthetic oligonucleotide complementary to the 5′ end of the longest cDNA sequence (nt 1–17, Figure 2) was labelled at its 5′ end with polynucleotide kinase and used to prime complementary strand synthesis on mRNA from interferon-treated and control Daudi (Figure 5A) and HeLa (data not presented) cells. In each case, on subsequent gel analysis, major and minor extension products (large and small arrows, Figure 5A) were observed with the interferon-induced but not the control RNA.

Table I. Exon/intron boundaries (A) and polyadenylation site (B) in 6-16 genomic DNA

A

DONOR SITE		ACCEPTOR SITE	
EXON-3′	5′ INTRON	3′	5′-EXON
Exon 1 38 ----AAG	GTAAAC----	CCCGTTCGCAG ----GGACCATCTCCCTCTCG	39 Exon 2 GT----
Exon 2 139 ----CAG	GTGAGA----	CTCTCCTGCAG ----CCACTACGCTTTGTCTG	140 Exon 3 GT----
Exon 3 217 ----GAG	GTGGGT----	TGTTCCTCCAG ----CCCTCACCTGCTCC	218 Exon 4 GA----
Exon 4 367 ----TCG	GTGAGT----	TTGCTTTACAG ----GACTCATCTGTGT	368 Exon 5 GG----
Consensus[a] ---- A/C AG	GT A/G AGT----	---(T/C)$_n$ N C/T AG	G----

B

PolyA

----AATAAACTTCACCCAGAAAACACTTTGTCC TGCT GTC------CTCGAG---

(760 at AATAAA; 775 at polyadenylation site, arrowed; +10; +100)

[a]Breathnach and Chambon (1981); Mount (1982) and Keller (1984).

(A) The GT–AG rule consensus sequences are boxed (solid lines) and the nt in accord with the more extensive consensus splice donor and acceptor site sequences are underlined. The nucleotides showing reasonable agreements (5 or 6 out of 7 residues) with the suggested (Keller, 1984; Reed and Maniatis, 1985) branch point consensus PyXPyTPuAPy are boxed (broken lines). The superscript numbers refer to the cDNA sequence in Figure 2. **(B)** The AATAAA consensus is underlined and the polyadenylation site arrowed. Superscript numbers refer to the cDNA sequence in Figure 2. Those downstream of the polyadenylation site are indicated by (+). The position of the *Xho*I site in the 3′-flanking region utilized in generating the 8.7-kb *Xho*I fragment subclone (text and Figure 4) is included (broken underline).

The major product (with both Daudi and HeLa mRNAs) was 54 nt in length, suggesting that the 5′ terminus of the mRNA lies 37 nt upstream of the 5′ end of the cDNA.

A genomic fragment containing DNA corresponding to the 5′ end of the mRNA was subcloned and sequenced (Figure 5B). As expected, it contained sequences in common with the first exon of the longest cDNA (compare the last 38 nt in Figure 5B with cDNA nt 1–38, Figure 2) as well as additional upstream sequences covering the proposed 5′ end of the mRNA. A TATA box sequence ATAAAA, identical to that used in the β-globin gene (Efstratiadis *et al.*, 1980), was found ~20 nt 5′ of the putative transcription initiation site, as is the case for the herpes simplex virus thymidine kinase (HSVtk) gene (McKnight and Kingsbury, 1982).

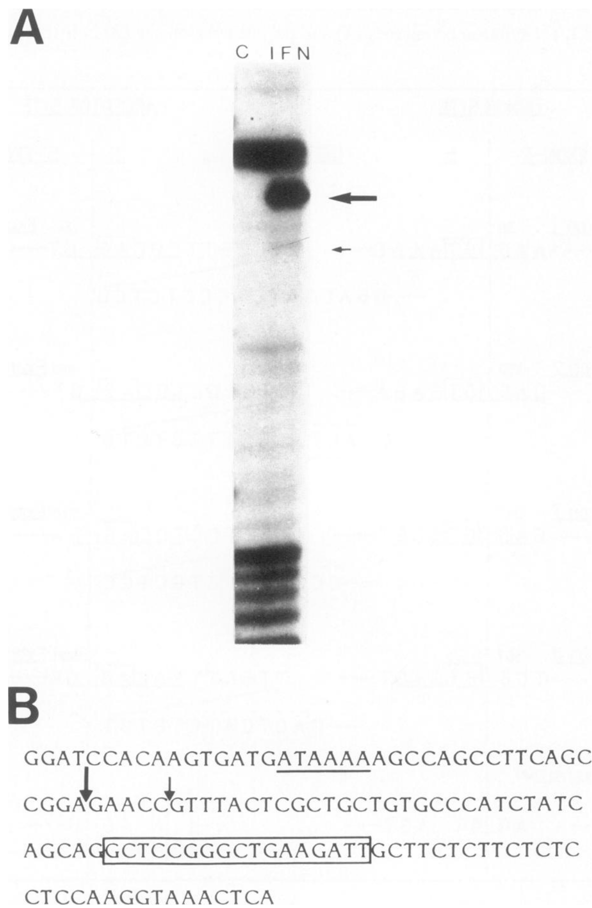

Fig. 5. **(A)** Primer extension of 6-16 mRNA. Extension using a 5′ ^{32}P-labelled synthetic oligodeoxynucleotide primer (see Materials and methods) was with mRNA from Daudi cells with (IFN) or without treatment with 100 IU/ml of human α-interferons for 16 h at 37°C. An autoradiograph of an electrophoretic analysis of the extended products on a sequencing gel is presented. The positions to which the major and minor 6-16-specific products unique to RNA from interferon-treated cells migrated, are indicated by large and small arrows respectively. The sizes of these products were determined by comparison with markers of known size in the same gel. **(B)** Sequence of the genomic DNA fragment containing the 5′ initiation site for transcription of the 6-16 mRNA. A 0.5-kb *Bam*HI genomic DNA fragment from cosmid 10.3 (Figure 4B) that specifically hybridized to the same 6-16-specific synthetic oligodeoxynucleotide primer as was used in the primer extension experiments was cloned into the *Bam*HI site of M13 mp11 and sequenced by the dideoxy chain termination method. The sequence for 118 nt downstream of the 5′ *Bam*HI site is presented. The position of the putative TATA box (ATAAAA) is underlined. The large and small arrows identify the positions corresponding to those at which the extended synthetic primers terminated in **(A)** above, indicative of major and minor transcription initiation sites respectively. The nt corresponding to the 5′ cDNA sequence (Figure 2) to which the (complementary) synthetic primer was made are boxed.

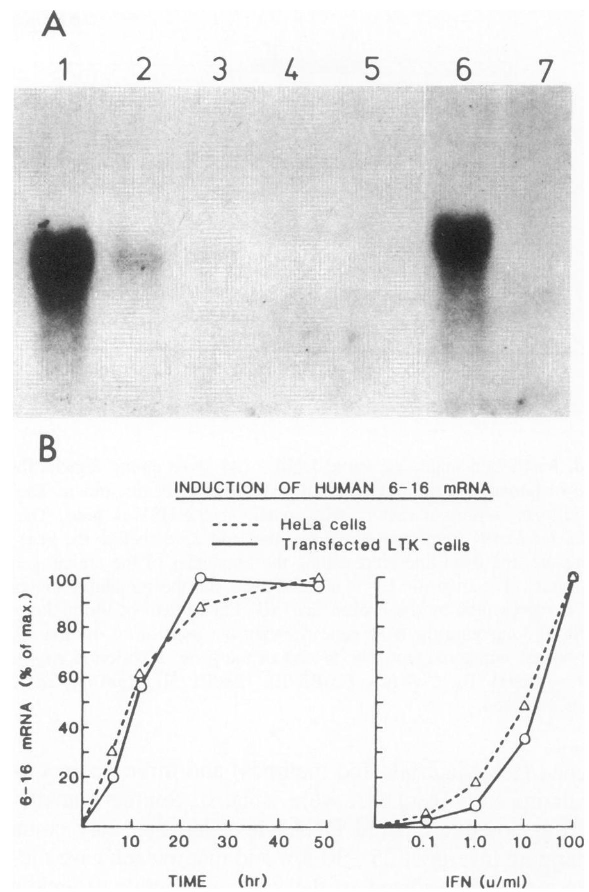

Fig. 6. Expression of human 6-16 mRNA in mouse cells transfected with cosmid 10.3. **(A)** Northern analysis (see Materials and methods) of cytoplasmic RNA from mouse Ltk⁻ cells transfected with **(tracks 1, 2** and **5)** cosmid 10.3 containing the entire 6-16 gene and 5′-flanking sequences or **(tracks 3** and **4)** cosmid 1.4 containing a 5′-truncated insert (Figure 4). The RNA (20 μg) was from cells grown to 50% confluency in monolayers and incubated for a further 17 h at 37°C in the presence **(tracks 1** and **3)** or absence **(tracks 2** and **4)** of 300 IU/ml of mouse α,β-interferons. **Tracks 5–7**: RNA from mouse Ltk⁻ cells transfected with cosmid 10.3 **(track 5)** or HeLa **(tracks 6** and **7)** cells grown in parallel, without **(track 7)** or with **(tracks 5** and **6)** similar exposure to 300 IU/ml of human α-interferons. **(B)** Kinetics and dose-response for the induction of human 6-16 mRNA in mouse cells: comparison with corresponding data for HeLa cells. Mouse cells transfected with cosmid 10.3 were treated with mouse α,β-interferon either at 300 IU/ml for the times indicated (left) or for 18 h with the indicated dose (right). HeLa cells were similarly treated with human α-interferons. Cytoplasmic RNAs were extracted, separated by electrophoresis and analysed by Northern blot hybridization as in **(A)**. The bands on the radiographs were scanned by densitometry. The values are presented as a percentage of the maximum level of 6-16 mRNA detected.

Interferon-inducible expression of human cosmid 10.3 in mouse cells

Cosmid clone 10.3 includes at least 20 kb of human DNA 5′ to the 6-16 gene. The overlapping cosmid clone 1.4 terminates in the first exon while extending substantially into the 3′-flanking region (Figure 4A). These cosmid DNAs were transfected into mouse Ltk⁻ cells and transfectants selected in HAT medium (the cosmid vector DNA contains the HSVtk gene). Mixed colonies (>60 clones) were used to guard against artefacts reflecting integration of the transfected DNA at particularly active or inactive sites in the genome in different cloned lines. RNA isolated from the transfected cells, with or without treatment with mouse α-, β-interferons, was analysed by Northern blot hybridization using a human 6-16 cDNA probe (Figure 6). There was no hybridization between the human 6-16 cDNA and mouse mRNA (e.g. tracks 3–5, Figure 6A). Cells transfected with cosmid 10.3 expressed the human 6-16 mRNA in response to mouse (track 1), but not human (track 5), interferons, whereas those transfected with cosmid 1.4 lacking the 5′ end of the gene (tracks 3 and 4) did not (Figure 6A). The kinetics and dose-response curves for induction were very similar to those obtained for 6-16 mRNA in human cells (Figure 6B). As in the human system, the induction of the human 6-16 mRNA in the transfected cell is a primary response (it occurs in the presence of sufficient cycloheximide to inhibit protein synthesis) and is restricted to the α- and β-interferons: mouse γ-interferons was without effect (data not shown).

DNA from a subclone of plasmid 10.3 containing only the

8.7 kb *Xho*I fragment (Figure 4B) yielded, on transfection, results similar to those obtained with the intact 10.3 cosmid DNA (A.C.G.Porter, data not shown). It can be concluded that cosmid 10.3 and this subclone contain the entire 6-16 gene and flanking DNA sequences essential for interferon induction.

Discussion

A human genomic cosmid corresponding to the interferon-inducible gene 6-16 has been isolated and characterized. The similarity in the inducibility of the human 6-16 mRNA in mouse cells transfected with this cosmid to that in HeLa cells (Figure 6B) strongly suggests that this cosmid contains an intact 6-16 gene and any flanking DNA necessary for a complete response to Type I interferons. A subclone of the cosmid containing only the 8.7-kb *Xho*I fragment of human DNA and hence only 2.25 kb of 5′ and 100 nt of 3′-flanking sequence (Figure 4B and Table IB) yielded similar results on transfection. It can be concluded that any essential DNA control elements must be present within this limited region. More detailed deletion analyses together with experiments to assess the inducibility of constructs in which putative control elements are placed 5′ of marker genes should permit a more precise definition of the sequences involved.

The induction of transfected human HLA genes by mouse interferon in mouse cells (Rosa *et al.*, 1983a; Yoshie *et al.*, 1984) has already established that the same or similar mechanisms must govern induction of these genes in the two species. It is curious that a mouse equivalent of the human 6-16 mRNA was not detectable even at low stringency in, for example, the transfection experiments (Figure 6A). This remains to be investigated further.

Comparison of appropriate genomic and cDNA sequences shows that the 6-16 gene contains five exons (Figure 4). The GT−AG rule is strictly adhered to in the sequences at the intron−exon boundaries, which are also in very good general agreement with the more extensive consensus sequences characteristic of these regions (Table IA; Breathnach and Chambon, 1981; Mount, 1982; Keller, 1984). As expected, the intron sequences upstream of the acceptor site are rich in cytidylic and thymidylic acid residues and do not contain AG doublets (Mount, 1982). In addition, in three cases (Table IA, broken line boxes, introns 2−4) the sequences around the A residue 19−23 nt 5′ of the acceptor site match reasonably well with a tentative branch site consensus sequence (Keller, 1984; Reed and Maniatis, 1985). In the case of the polyadenylation site (Table IB), the 'dominant' consensus sequence AATAAA is adhered to and the additional downstream CA$^{C}_{T}$TG (Berget, 1984), AAA and GT (McLauchlan *et al.*, 1985) cluster motifs, or close approximations to them, are all present, although not in an ideal pattern (reviewed by Birnstiel *et al.*, 1985). The one obvious divergence is with the terminal C(A) consensus (for the point of addition of the poly(A), Birnstiel *et al.*, 1985), which is not observed.

The exact position of the initiation of transcription remains to be rigorously established. The results of the primer extension experiments (Figure 5) place the 5′ end of the mRNA 37 nt upstream of the 5′ end of the cDNA sequence (Figure 2), in reasonable agreement with the apparent size of the mRNA allowing for the poly(A) tail (Kelly *et al.*, 1985; Figure 1). The presence in the genomic DNA (Figure 5B) of a TATA box (identical to that used for the β-globin gene) ~20 nt 5′ of the initiation site (indicated by the primer extension experiments), together with the results of initial cell-free transcription assays (Y.Chernjovsky, unpublished data), suggest that the proposed initiation site is indeed correct. Although the presence of an additional very short 5′ exon cannot be excluded absolutely, it seems highly unlikely. Were such an exon to exist, the activity of the transfected 8.7-kb *Xho*I fragment establishes that it would have to lie within the 2.25 kb of 5′-flanking sequence and should be resolved by the further analysis of this fragment.

Evidence that the longest open reading frame deduced from the cDNA sequence (Figure 2) is indeed that employed comes from recent work of Evans, Secher and Milstein (personal communication) in which cDNA clones analogous to those reported here were isolated using antibody to the interferon-inducible membrane protein P16 (Burrone and Milstein, 1982). Assuming that P16 is indeed the product of the 6-16 gene, some index of its function may now be obtained by, for example, studying the effects of its high level expression from suitable constructs in cells which have not been treated with interferon.

Materials and methods

Interferons

Wellferon, a highly purified mixture of human α-interferons (> 10^8 IU/mg protein, Allen *et al.*, 1982) was supplied by Dr K.Fantes, Wellcome Research Laboratories. Natural human β-IFN (10^7 IU/mg protein, a gift of Dr John Davies, Searle Research and Development Division, High Wycombe) was from superinduced cultures of human diploid fibroblasts (McCullagh *et al.*, 1983). Non-glycosylated human γ-IFN (> 10^7 IU/mg protein) was from Biogen Ltd. Murine α,β-interferon, a natural mixture of virally induced interferons ($1-5 \times 10^7$ IU/mg protein) was kindly supplied by Dr S.L.Mowshowitz, Mount Sinai Hospital, NY. Recombinant murine γ-interferon ($1-2 \times 10^7$ IU/mg protein) was from Genentech.

Isolation and sequencing of cDNA clones

The original 6-16 cDNA was obtained by differential screening of a λgt10 cDNA library prepared with mRNA from α-interferon-induced T98G neuroblastoma cells (Friedman *et al.*, 1984). This clone was used to isolate several longer cDNAs from a λgt10 library constructed by M.McMahon and R.L.Friedman with mRNA from α-interferon-treated Daudi cells. The sequence of each of the cDNA inserts was determined for both strands by sequencing smaller DNA fragments (generated using the restriction enzymes *Alu*I, *Hin*fI or *Sau*3A and subcloned into M13) by the dideoxy chain termination method of Sanger (1981).

Isolation of genomic clones from a human DNA cosmid library

A library of human placental DNA fragments (>25 kb) cloned into the *Cla*I restriction site of the cosmid vector pHC79-2 cos/tk (Hohn and Collins, 1980; Lindenmaier *et al.*, 1982) was kindly supplied by Dr John Collins, Gesellschaft für Biotechnologische Forschung GmBH, Braunschweig-Stockheim. The library, in the form of a cosmid particle suspension, contained approximately 300 000 clones in 12 pools. *Escherichia coli* HB101, prepared for transformation by growth overnight in NZYCM medium (Maniatis *et al.*, 1982) plus 0.4% maltose, were concentrated 2.5-fold by resuspension in 10 mM $MgSO_4$. Aliquots (100 μl) were infected with 1.5×10^5 p.f.u. of one of the cosmid pools and incubated for 20 min at 37°C. NZYCM medium (0.5 ml) was added and each pool incubated for a further 45 min at 37°C. The 12 pools combined in pairs were plated onto six Biodyne membranes (Pall Corp) on NZYCM agar plates containing ampicillin (50 μg/ml) and incubated overnight at 37°C. Replica lifts were made onto a further six membranes which were incubated for 6−7 h at 37°C on NZYCM agar ampicillin plates as above until colonies became visible. The master plates were stored at 4°C. The replica filters were transferred to six agar plates containing chloramphenicol (150 μg/ml) and incubated overnight at 37°C to amplify the number of cosmids per cell. The filters were removed, the *E. coli* lysed in alkali and the DNA bound by baking at 80°C for 2 h (Maniatis *et al.*, 1982). Hybridization and prehybridization conditions were identical to those used for Southern analysis (see below) except that the prehybridization was preceded by a prewash in 50 mM Tris-HCl, pH 8.0, 1 M NaCl, 1 mM EDTA, 0.1% SDS for 2 h at 42°C. The filters were probed with a nick-translated 6-16 cDNA insert (700 nt, ~10^8 c.p.m./μg) obtained after *Eco*RI digestion of the replicative (double stranded) form of an M13 clone. Positive colonies were detectable after autoradiography at −80°C for 1−2 days. Using the autoradiographs as templates, positive colonies were picked from the master plates. After two further rounds of screening, seven colonies which hybridized to the 6-16 probe were isolated. DNA minipreps (Maniatis *et al.*, 1982) were prepared from 10 ml overnight cultures (in 50 μg/ml ampicillin) of each of the purified colonies. Each DNA preparation was dissolved in 200 μl of 10 mM Tris, pH 7.6, 1 mM EDTA and stored at −20°C. Aliquots (20 μl) were incubated with *Eco*RI and the DNA separated by electrophoresis through 0.8% agarose gels. Ethidium bromide staining and subsequent Southern analysis revealed that the seven positive colonies represented three distinct overlapping genomic clones.

Sequence analysis of the genome clones for 6-16

The regions of human genomic cosmid DNA which contained mRNA coding sequence were digested with the four-cutter restriction enzyme *AluI*, *Sau3A* and *HinfI* to yield overlapping fragments which were subcloned into the *SmaI* site of the M13 mp18 vector (Norrander *et al.*, 1983). DNA was prepared and subclones containing exon sequences were identified by dot blot hybridization against a nick-translated 700 nt fragment of 6-16 cDNA. These DNA subclones were sequenced as above. Comparison of the genomic with the cDNA sequences enabled each of the exon–intron junctions to be mapped.

Preparation of RNA and DNA from HeLa and L-cells

Cells were grown in monolayer culture in Dulbecco's modified Eagle's medium supplemented with 10% heat-inactivated calf serum. When the cells had reached approximately 50% confluency interferon was added as indicated. For total RNA, the cells were lysed by adding 4 M guanidinium thiocyanate, 0.5% sodium *N*-lauryl sarcosine (SDS), 25 mM sodium citrate, pH 7.0. For cytoplasmic RNA the cells were first lysed in 0.5% NP40 and the RNA was extracted from the cytoplasmic fraction as above. The RNA was pelleted by centrifugation through a 5.7 M CsCl gradient, precipitated twice with ethanol and stored in water at −80°C. The accumulation of mRNA in the absence of protein synthesis was investigated by adding cycloheximide (50 μg/ml) to the growth medium 30 min before interferon treatment. Poly(A)$^+$ RNA was prepared from total RNA by standard techniques as was genomic DNA (Maniatis *et al.*, 1982).

Northern analysis of interferon-inducible mRNAs

Aliquots of RNA (20 μg) were glyoxalated and fractionated through 1.4% agarose gels. The RNA was transferred to Biodyne nylon filters (Pall Corp.) which were baked at 80°C for 2 h. These were hybridized with ^{32}P-labelled, single-stranded M13 probes in 50% formamide, 5 × SSPE (900 mM NaCl, 5 mM EDTA, 50 mM sodium phosphate, pH 7.4), 5 × Denhardt's solution, 0.2% SDS and 100 μg/ml denatured salmon sperm DNA for 18 h at 42°C. The filters were washed for 1 h in 2 × SSPE, 0.1% SDS at 50°C and then for 1 h in 0.1 × SSPE, 0.1% SDS at 50°C. Low stringency washes were carried out in 6 × SSPE, 0.1% SDS at room temperature. Kodak XAR-5 film was used for autoradiography. The extent of RNA induction was established by scanning the appropriate band on autoradiographs using a Joyce-Loebl Chromoscan 3 densitometer. Different exposures were scanned to ensure that there was a linear relationship between the amount of radioactivity on the filter and the resultant darkening of the film.

Southern analysis of genomic and cosmid DNA

Aliquots of genomic (20 μg) or cosmid (~2 μg) DNA were digested with restriction enzymes and the fragments fractionated through 0.8% agarose gels. The DNA was transferred to Biodyne membrane filters which were baked at 80°C for 2 h prior to hybridization with nick-translated DNA (10^8 c.p.m./μg) or ^{32}P-labelled single-stranded M13 DNA probes in 50% formamide, 5 × SSPE, 1 × Denhardt's solution, 0.2% SDS, 100 μg/ml denatured salmon sperm DNA for 18 h at 42°C. The filters were washed and autoradiographed as described for Northern analysis.

Primer extension

An oligodeoxynucleotide primer 5'-AATCTTCAGCCCGGAGC, complementary to the 5' residues (17–1) of the 6-16 cDNA sequence shown in Figure 2, was purchased from Dr Stephen Minter, University of Manchester Institute of Science and Technology. The extension reactions were essentially as described by Houghton *et al.* (1980). Briefly, 30 pmol of primer, 5'-end-labelled with ^{32}P using polynucleotide kinase, was lyophilized with 1.2–1.5 μg of total poly(A)$^+$ RNA, resuspended in 4 μl of 0.4 M KCl, heated to 100°C for 1 min and hybridized for 1 h at 37°C. The hybridization mixture was incubated with 30 or 40 units of MLV or AMV reverse transcriptase (B.R.L. and Anglian Biotech respectively) in 50 mM Tris-HCl, pH 8, 0.5 mM (each) dNTPs, 5 mM dithiothreitol, 4 mM $MgCl_2$ (final volume 25 μl) for 2 h at 37°C. The RNA was extracted with phenol, precipitated with ethanol plus 10 μg of carrier *E. coli* tRNA, resuspended in 2 μl 0.3 M NaOH plus 3 μl of 80% formamide in 50 mM Pipes buffer pH 7; 0.1% bromophenol blue; 0.1% xylene cyanol, heated to 100°C for 5 min and separated by electrophoresis through a 20% acrylamide/urea sequencing gel (Maniatis *et al.*, 1982). Autoradiography with Fuji X-ray film and intensifier screen was overnight at −70°C.

Transfections

Mouse Ltk$^-$ cells were transfected by a calcium phosphate procedure (Parker and Stark, 1979). Dulbecco's modified Eagle's medium was used throughout. Either the tk$^+$ cosmid DNA alone or a mixture of tk$^-$ test plasmid DNA with tk$^+$. pAGO DNA (Colbere-Garapin *et al.*, 1979) at a molar ratio of 10:1 (tk$^-$:tk$^+$) was used. The DNA (20 μg) was first precipitated with ethanol and then resuspended in 40 μl of 10 mM Tris-HCl, pH 7.2; 1 mM EDTA and 440 μl 42 mM Hepes buffer, pH 7; 137 mM NaCl; 10 mM KCl; 1.6 mM NaH_2PO_4. $CaCl_2$ (880 μl of 260 mM) was then added dropwise with vigorous mixing to form a fine precipitate and after 30 min at room temperature the suspension was transferred to cells which had been seeded at 10^6 per 9-cm dish, grown overnight and replenished with 7 ml of medium. After 8–16 h of incubation at 37°C with the DNA, the cells were washed and treated for 1 min with medium containing 25% (v/v) glycerol (Parker and Stark, 1979) and for 12 h in medium containing 5 mM sodium butyrate (Gorman and Howard, 1983). They were then grown in medium containing 0.1 mM hypoxanthine, 0.4 μM aminopterin and 16 μM thymidine for 1–2 weeks until stable tk$^+$ colonies were visible (60–1000 per dish).

Acknowledgements

We are very grateful to John Collins for the gift of the cosmid library and to Tom Evans, David Secher and Cesar Milstein for communicating their results to us prior to publication. Y.C. was the recipient of a long-term EMBO postdoctoral fellowship.

References

Allen,G., Fantes,K.H., Burke,D.C. and Morser,J. (1982) *J. Gen. Virol.*, **63**, 207–212.

Benech,P., Mory,Y., Revel,M. and Chebath,J. (1985) *EMBO J.*, **4**, 2249–2256.

Berget,S.M. (1984) *Nature*, **309**, 179–182.

Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349–359.

Burrone,O.R. and Milstein,C. (1982) *EMBO J.*, **1**, 345–349.

Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.

Colbere-Garapin,F., Chousterman,S., Horodniceanu,F., Kourilsky,P. and Garapin,A.-C. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3755–3759.

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. and Proudfoot,N.J. (1980) *Cell*, **21**, 653–668.

Friedman,R.L. and Stark,G.R. (1985) *Nature*, **314**, 637–639.

Friedman,R.L., Manly,S.P., McMahon,M., Kerr,I.M. and Stark,G.R. (1984) *Cell*, **38**, 745–755.

Gorman,C.M. and Howard,B.H. (1983) *Nucleic Acids Res.*, **11**, 7631–7648.

Hohn,B. and Collins,J. (1980) *Gene*, **11**, 291–298.

Houghton,M., Stewart,A.G., Doel,S.M., Emtage,J.S., Eaton,M.A.W., Smith, J.C., Patel,T.P., Lewis,H.H., Porter,A.G., Birch,J.R., Cartwright,T. and Carey,N.H. (1980) *Nucleic Acids Res.*, **8**, 1913–1931.

Keller,W. (1984) *Cell*, **39**, 423–425.

Kelly,J.M., Gilbert,C.S., Stark,G.R. and Kerr,I.M. (1985) *Eur. J. Biochem.*, **153**, 367–371.

Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857–872.

Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.

Larner,A.C., Jonak,G., Cheng,Y-S.E., Korant,B., Knight,E. and Darnell,J.E., Jr (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6733–6737.

Lindenmaier,W., Hauser,H., Greiser de Wilke,I. and Schütz,G. (1982) *Nucleic Acids Res.*, **10**, 1243–1256.

Luster,A.D., Unkeless,J.C. and Revetch,J.V. (1985) *Nature*, **315**, 672–676.

McCullagh,K.G., Davies,J.A., Sim,I.S., Dawson,K.M., O'Neill,G.J., Doel, S.M., Catlin,G.H. and Houghton,M. (1983) *J. Interferon Res.*, **3**, 97–111.

McKnight,S.L. and Kingsbury,R. (1982) *Science*, **217**, 316–324.

McLauchlan,J., Gaffney,D., Whitton,J.L. and Clements,J.B. (1985) *Nucleic Acids Res.*, **13**, 1347–1368.

McMahon,M., Stark,G.R. and Kerr,I.M. (1986) *J. Virol.*, **57**, 362–366.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.

Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.

Norrander,J., Kempe,T. and Messing,J. (1983) *Gene*, **26**, 101–106.

Parker,B.A. and Stark,G.R. (1979) *J. Virol.*, **31**, 360–369.

Reed,R. and Maniatis,T. (1985) *Cell*, **41**, 95–105.

Rosa,F., Le Bouteiller,P.P., Abadie,A., Mishal,Z., Lemonnier,F.A., Bourrel,D Lamotte,M., Kalil,J., Jordan,B. and Fellous,M. (1983a) *Eur. J. Immunol.*, **13**, 495–499.

Rosa,F., Hatat,D., Abadie,A., Wallach,D., Revel,M. and Fellous,M. (1983b) *EMBO J.*, **2**, 1585–1589.

Samanta,H., Chao,H., Engel,D., Tominaga,S., Vijayasarathy,s. and Lengyel,P. (1984) *Antiviral Res.*, **1**, 22.

Sanger,F. (1981) *Science*, **214**, 1205–1210.

Sharp,P.A. (1983) *Nature*, **301**, 471–472.

Staeheli,P., Haller,O., Boll,W., Lindenmann,J. and Weissmann,N.C. (1986) *Cell*, **44**, 147–158.

Strunk,R.C., Cole,F.S., Perlmutter,D.H. and Colten,H.R. (1985) *J. Biol. Chem.*, **260**, 15 280–15 285.

von Heijne,G. (1984) *J. Mol. Biol.*, **173**, 243–251.

von Heijne,G. (1985) *J. Mol. Biol.*, **184**, 99–105.

Yoshie,O., Schmidt,H., Lengyel,P., Reddy,E.S.P., Morgan,W.R. and Weissman,S.M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 649–653.

Received on 20 March 1986