The EMBO Journal vol.7 no.1 pp.85 – 92, 1988

Interferon response element of the human gene 6-16

Andrew C.G.Porter, Yuti Chernajovsky[1], Trevor C.Dale, 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: Department of Immunology, M.D.Anderson Hospital and Tumor Institute, University of Texas System Cancer Center, Houston, TX 77030, USA

Communicated by I.M.Kerr

1046 base-pairs (bp) of genomic DNA spanning the first exon of the human α/β-interferon(IFN)-inducible gene 6-16 have been analysed for their role in induction. The whole gene or 5′-flanking deletion derivatives of it were assayed for inducibility in populations of stably transfected mouse cells. 5′-Flanking DNA fragments were assayed for their ability to confer inducibility on a reporter gene in stably and transiently transfected mouse and human cells. The data suggest that a 39 bp sequence is sufficient to confer transcriptional inducibility and can account in large part for the response of 6-16. Two copies of this sequence, one of which contains a dinucleotide insert, are located in tandem 88 bp upstream of the 6-16 transcriptional initiation site. For at least one of the repeat units the 5′ limit of a subregion required for induction lies in the sequence GGGAAAAT. The motif GGAAA occurs in several well characterized enhancers. Furthermore, one residue 3′ of the GGAAA there is a second motif, TGAAACT, which is conserved in the regulatory regions of other IFN-induced genes. In gel retardation assays the oligonucleotide GGGAAAATGAAACT competes with the repeat element for binding to IFN-modulated protein(s) but a mutated oligonucleotide, GGGAAAATGACACT does not. These results identify an α/β IFN response element partially homologous to those described previously for the genes of the MHC complexes.

Key words: enhancer/gene/human/interferon-inducible/transcription

Introduction

The genes coding for a number of interferon (IFN)-inducible proteins have been isolated in order to study their regulation and function (Chebath *et al.*, 1983; Friedman *et al.*, 1984; Larner *et al.*, 1984; Reich *et al.*, 1987; Merlin *et al.*, 1983; Staeheli *et al.*, 1986; reviewed by Revel and Chebath, 1986). Although evidence exists for both transcriptional and post-transcriptional control (e.g. Friedman *et al.*, 1984), the nucleotide sequences that mediate these controls have not been characterized extensively. A knowledge of the sequences governing transcriptional control and the cellular factors with which they interact would be a first step towards understanding the signalling mechanisms that operate between the binding of the IFNs to their cell-surface receptors and the induction of gene expression.

A potential IFN response sequence (IRS) was noted on the basis of conserved sequences in the 5′-flanking regions of human class I and class II HLA and MTIIa genes (Friedman and Stark, 1985). Similar sequences have since been found in association with the genes for complement Factor B (Wu *et al.*, 1987), β2-microglobulin and the mouse H-2 complex (Kimura *et al.*, 1986). In the case of the H-2 complex, the IRS has been shown to be functionally involved in the induction by type I (α/β) IFNs (Israel *et al.*, 1986, 1987; Korber *et al.*, 1987). A distinct response element may be involved in the induction of class II HLAs by type II (γ) IFN (Boss and Strominger, 1986). In none of these cases, however, is the extent of induction particularly large. Accordingly, we have chosen to study the human gene 6-16 which is strongly and preferentially induced by type I IFNs (Friedman *et al.*, 1984).

The 6-16 gene can code for a 12 kd hydrophobic protein (Kelly *et al.*, 1986). Preferential induction by type I IFNs has been observed in several human cell-types (Kelly *et al.*, 1985). In common with other IFN-inducible genes, the induction of 6-16 does not usually require new protein synthesis, indicating that pre-existing cellular proteins are involved (Friedman *et al.*, 1984). Exceptionally, protein synthesis is required for induction of 6-16 in HeLa M cells (Kusari and Sen, 1987). A cosmid (10.3) carrying the entire 6-16 gene as well as ~20 kbp of 5′-flanking sequences is regulated normally by mouse IFN in populations of stably transfected mouse cells (Kelly *et al.*, 1986). Here we have determined at what point 5′ deletions extending towards the first exon cause an impairment or loss of this inducibility. In addition, sequences from the 5′ end of the gene have been assayed for their ability to confer inducibility on reporter genes and to bind proteins in gel retardation assays. Recently Reich *et al.* (1987) and D.Levy, N.Reich and J.Darnell (personal communication) have identified for the ISG15 and 54 genes a region of homology (16 out of 20 bases, at approximately −90 to −110) within the 5′-flanking sequences shown by deletion and construct analysis to govern the IFN response. Comparison with upstream sequences of 6-16 reveals a similar homology, thus strengthening the conclusion that an IFN response element has been identified.

Results

IFN-inducibility of subclones of cosmid 10.3

An 8.7 kb DNA fragment containing the entire 6-16 gene plus 2.3 kb of 5′- and 100 bp of 3′-flanking sequence was subcloned from the cosmid 10.3 into the vector pGEM2 to generate the plasmid shown in Figure 1A (Kelly *et al.*, 1986). Populations of Ltk^- cells stably transfected with this plasmid, a derivative containing only 603 bp of 5′-flanking sequences (Figure 1Ab) or the cosmid 10.3, were grown

in the presence or absence of IFN and assayed for human 6-16 mRNA (Figure 1B). Although there was some variation between populations in the absolute amounts of 6-16 mRNA, the response to IFN was conserved in each population. Sequences required for inducibility were therefore concluded to be 3′ of the *Bgl*II site at −603. [Numbers refer to nucleotide positions 5′ (negative) or 3′ (positive) of the transcriptional initiation site for 6-16.]

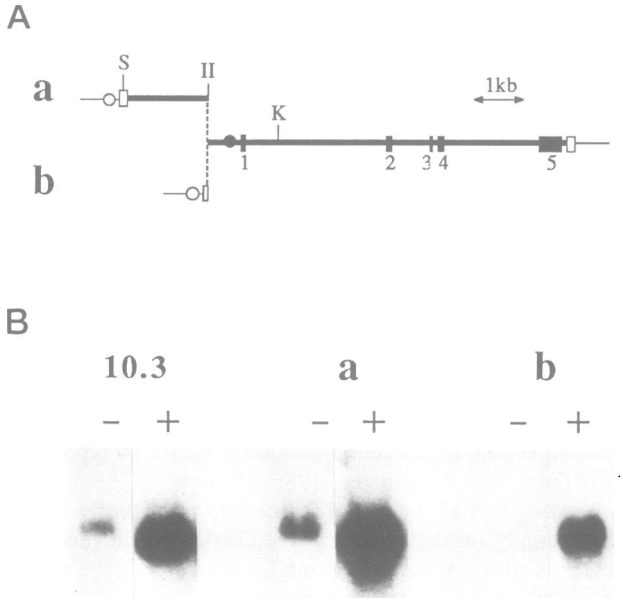


Fig. 1. (**A**) Restriction map of two subclones from cosmid 10.3 containing the whole 6-16 gene. Constructs **a** and **b** are identical except for the 1.7 kb *Sma*I/*Bgl*II fragment which is deleted in **b**. Human 6-16 sequences (thick horizontal lines) including exons (numbered solid boxes) are drawn to scale. The position of the tandem duplication (see text) is shown as a solid circle. Vector (pGEM2) sequences (thin horizontal lines) including the polylinkers (open boxes) and the bacteriophage SP6 promoter (open circle) are not drawn to scale. Unique sites for the enzymes *Bgl*II (II), *Kpn* (K) and *Sma*I (S) are indicated. (**B**) Northern analysis of 6-16 mRNA from L*tk*$^-$ cells stably transfected with cosmid 10.3 or constructs **a** or **b**. Transfectants were grown to ~50% confluence and then for a further 22−24 h in the presence (+) or continued absence (−) of IFN (300 IU/ml). Preparation and analysis of RNA was as described (Materials and methods). Under the hybridization conditions used, an endogenous 6-16 mRNA is not detectable in L*tk*$^-$ cells (Kelly *et al.*, 1986).

Nucleotide sequences surrounding exon 1

The nucleotide sequence from the *Bgl*II site at −603 to the *Bam*HI site at +437 is shown in Figure 2. A notable feature of this sequence is the presence of a tandem duplication at position −89 to −168. The first (−168 to −128, 41 bp) and second (−127 to −89, 39 bp) copies of the repeat unit are identical except for the dinucleotide CT (−139, −138) which is present in the first but not in the second copy. A sequence homologous to the CCAAT box of the herpes simplex virus thymidine kinase (HSV*tk*) gene (Graves *et al.*, 1986) is located (−63 to −67, opposite strand) between the tandem duplication and the TATA box. Additional homologies are discussed below with reference to Table I and Figure 7.

Inducibility of 6-16 CAT constructs

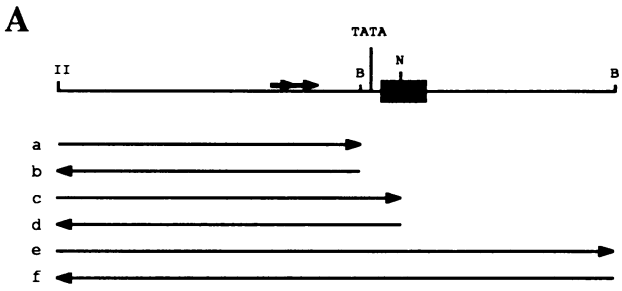
DNA fragments extending from the *Bgl*II site (−603) to sites upstream (−39) within (+42) or beyond (+437) the first exon (Figure 3A) were assayed for their ability to confer inducibility on the gene for the bacterial enzyme chloramphenicol acetyl transferase (CAT) when cloned upstream of the CAT gene in the vectors pSVOcat or $pA_{10}cat_2$ (Gorman *et al.*, 1982; Laimins *et al.*, 1982; Rosenthal *et al.*, 1983). Both vectors lack enhancers and express minimal CAT activity. In pSVOcat there are no promoter sequences upstream of the CAT gene whereas in $pA_{10}cat_2$ a 200 bp region of the SV40 early promoter, including the TATA box, is located between the cloning site and the CAT gene. Therefore, pSVOcat and $pA_{10}cat_2$ were used to test fragments with or without the 6-16 TATA box, respectively. The results for the IFN inducibility of CAT activity in populations of stably transfected L*tk*$^-$ cells are shown in Figure 3B. Similar inducibilities were obtained in transiently transfected L*tk*$^-$ cells and in stably or transiently transfected HeLa cells (data not shown). In all cases, CAT activity accumulated in response to IFN with kinetics similar to the kinetics of accumulation of 6-16 mRNA (Friedman *et al.*, 1984).

Inducibility of the construct containing only 5′-flanking 6-16 DNA (construct a, Figure 3A and B) indicates that

```
-604  AGATCTATCATGAT GGCCACATGA ACACAGGCTT CACTGGTCTT ACCATATGCC CATGACCCAG AAGCAGCCAG CCTGAGAGAA CAATGGAACA AAGTGACTTA   -501
-500      ATTTCCAAAG TTCTGGGGGT TTATAAACAG CATATGGTAT GTCACATGAC TACTGTGGGA TTGGAAAAAG ATCATGTAAA TAGAGCACTT TGCACAGACC   -401
-400      CTGGCATGCA GCAAGTGCTC AATAAATGAT AGGTGTTGTT TACTAATTGG ACTGAATGGT GAAAGGCCTG TGTGCCCCAG GGGGAGCTGG TGATCAGGCT   -301
-300      TCACTAAGCC CAGTATGGCC GTGGCTCTCA TCATAGTGTG ACATGCTTTG AATACCCTTA GCGGCTCCAA AAGTCCTCAG CTTGAATGCA TTTTTCTGCC   -201
-200      AGGCAGGCAG GCACACAAAT GTTCCGCTGG GCGGAGCTGG GAGAGAGGGG AAAATGAAAC TCTGCAGAGT GCAGGAGCTG GGAGAGAGGG GAAAATGAAA   -101
-100      CTGCAGAGTG CAGAAATAGA AACTCCGACA GGGATTGGCT GCCTAGGGTG AGACGTGGGA GGATCCACAA GTGATGATAA AAAGCCAGCC TTCAGCCGGA     -1
   1      GAACCGTTTA CTCGCTGCTG TGCCCATCTA TCAGCAGGCT CCGGGCTGAA GATTGCTTCT CTTCTCTCCT CCAAGGTAAA CTCAGGAGCT TATGAAGTGT    100
+101      GGGCATTCAA GCTGCCACCC TCTGCCAGGC TGCCTGTCTG CCTGTAAATC TCATGTTCTG AGAGCCAGGA GGCCCCTTCT CCTGGGAGGC AGCACTCCTG    200
+201      GGTCCCTTTT AGTGCTCTGG GCTGGGACTT GTCTAAGAGG ATGGGTTGGA GATTTTTAGG GAGATGGGAT GCAAAACCCC AAGTGGCATG AGACCCAGCT    300
+301      TACAGGTGCA ATATCAGCGA TCTGTGGCCT TAACACTGTC ACCTCTTGGA GCCTTAATTA CTTCCTCTGT AAAAGGAAAG TTAAGTTGCC TTTGCTGCTC    400
+401      AAAGGACTGG ATGATTTTAG AATCCCCTTT ATTATAGGAT CC
```

Fig. 2. Nucleotide sequence of the 5′ end of the 6-16 gene. Sequence from the *Bgl*II (position −603) to the *Bam*HI (position +437) sites is shown. Exon 1 and the TATA box, the positions of which were determined previously (Kelly *et al.*, 1986), are outlined. The positions of the repeat units are indicated by arrows. The sequence motif GGGAAAAT, present in each repeat and shown to be required for induction, at least for the second repeat, are overlined. Homology to the HSV*tk* CCAAT box is underlined (−63 to −67, opposite strand). ▼: the 3′ ends of various deletions generated by nuclease *Bal*31 digestion (see text).▽: the *Pst*I sites used to generate a 6-16 gene lacking the second repeat unit (Materials and methods) are shown.

regulation occurs primarily at the level of transcription. Constructs analogous to a and b (Figure 3A), but with their inserts cloned into a site (*Bam*HI) 3′ of the CAT gene of $pA_{10}cat_2$, also showed inducible CAT activity (data not shown). The ability of inserts to confer inducibility when inverted (e.g. constructs b,d and especially f, Figure 3) and with the repeat units at different positions relative to the CAT gene is consistent with an inducible enhancer mechanism (Maniatis *et al.*, 1987), although in two cases (constructs d and f, Figure 3) it is not possible to predict the TATA box and transcriptional initiation site involved. The greater absolute amount of CAT activity for constructs c and e, relative to construct a, may reflect an increased stability or translational efficiency of the longer CAT transcripts. The greater relative inducibility of constructs c and e may reflect a contribution to the induction mechanism from their additional 6-16 sequences (i.e. −39 to +437) despite the fact that these sequences alone do not confer inducibility (data not shown). However, it cannot be excluded that the observed differences in inducibility reflect unpredictable differences between constructs in the proportion of constitutive CAT activity resulting, for example, from incorrectly initiated transcripts. Whatever the explanation, the data clearly establish that inducibility can be conferred on a marker gene by sequences originating entirely from the 5′-flanking sequence between −603 and −39 (a, Figure 3A and B). A deletion analysis of this region was therefore carried out.



B

Insert	Vector	IFN	CAT Activity 12h	24h	48h
-	$pA_{10}cat_2$	-	0.36	0.35	0.35
		+	0.38	0.36	0.28
a	"	-	1.6	2.0	1.6
		+	4.1	7.7	9.1
b	"	-	0.89	1.1	0.72
		+	1.1	1.3	1.6
-	pSVOcat	-	2.6	2.7	2.3
		+	2.4	2.3	1.6
c	"	-	5.5	4.2	3.8
		+	18	29	35
d	"	-	0.07	0.10	0.06
		+	0.11	0.14	0.15
e	"	-	26	27	21
		+	210	330	360
f	"	-	0.05	0.05	0.06
		+	0.22	0.43	0.49

Fig. 3. Design and analysis of 6-16-CAT constructs. (**A**) Map of the 1046-bp *Bgl*II/*Bam*HI fragment of 6-16 (Figure 2, for sequence) including the repeat units (thick arrows) and the first exon (solid box). The position of the TATA box and sites for the enzymes *Bgl*II (II), *Bam*HI (B) and *Nci*I (N) are shown. These enzymes were used to generate fragments (thin arrows) that were cloned into the vectors $pA_{10}cat_2$ or pSVOcat (Materials and methods) in both orientations with the tandem duplication directed towards (**a,c,e**) or away from (**b,d,f**) the CAT gene. (**B**) CAT activity in lysates from Ltk^- cells stably transfected with constructs shown in (**A**) and corresponding vector controls. Treatment with IFN was for the times shown. Each number represents the average activity (Materials and methods) for duplicate wells of cells (duplicates agreed to within 10%).

5′-Deletion analysis of the whole 6-16 gene

A series of plasmids similar to the one shown in Figure 1Ab, but with deletions of varying lengths (▼, Figure 2) starting at the *Bgl*II site and extending towards exon 1, was generated (Materials and methods). Populations of Ltk^- cells stably

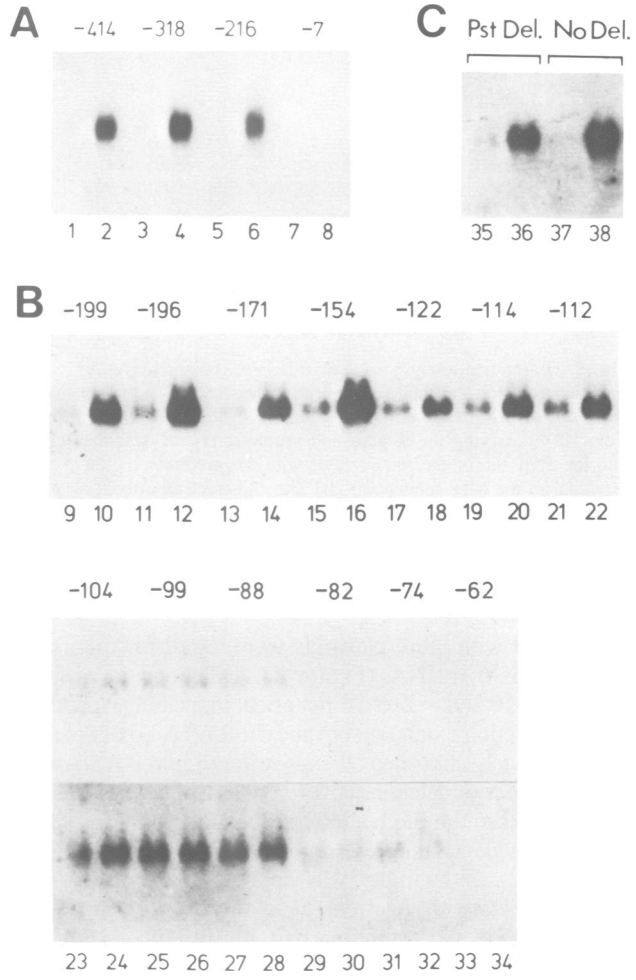


Fig. 4. Northern analysis of RNA from Ltk^- cells stably transfected with 5′ deletion constructs of the 6-16 gene. All of the constructs were generated (Materials and methods) from the plasmid shown in Figure 1Aa. An analysis of the RNA from cells transfected with this plasmid is included for comparison (**C**, **lanes 37** and **38**). (**A**) and (**B**) *Bal*31 deletion constructs: the limits of the deletions relative to the 6-16 transcriptional initiation site are indicated by numbers above the lanes and by ▼ in Figure 2. (**C**) *Pst*I deletion construct: transfection was with a construct (Pst Del, **lanes 35** and **36**) identical to that in Figure 1Aa except that the 39 bp fragment between the *Pst*I sites (▽, Figure 2) was deleted (the effect of this deletion being equivalent to the removal of the downstream 39 bp repeat). RNA was isolated from cells grown for 22−24 h in the presence (even numbered lanes) or absence (odd numbered lanes) of IFN (300 IU/ml) and subjected to Northern analysis as described (Materials and methods). Two exposures of the autoradiogram for **lanes 23−34** are shown; the duration of the upper (shorter) exposure was the same as that shown for **lanes 9−22**. Growth, IFN-induction and Northern analysis of transfectants were repeated at least twice for all of the deletion constructs with no significant changes in inducibility.

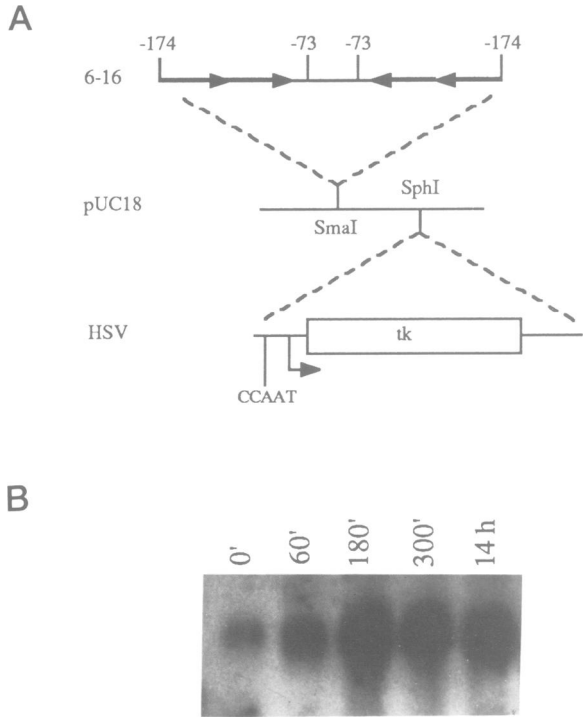


Fig. 5. Design and analysis of a 6-16-HSV*tk* construct. (**A**) A fragment carrying the tandem duplication was cloned as shown into the *Sma*I site of pUC18 (Materials and methods). A 1.8-kb *Pvu*II fragment from HSV, carrying the *tk* gene and promoter region, was then cloned into the *Sph*I site of the same vector with the promoter region proximal to the 6-16 sequences. (**B**) The construct described in (A) was cotransfected with pSV2hygro into L*tk*$^-$ cells and a stable population of hygromycin-resistant transfectants was selected. This population was treated with IFN (300 IU/ml) for the times indicated before the isolation and Northern analysis of RNA using a *tk* probe (Materials and methods).

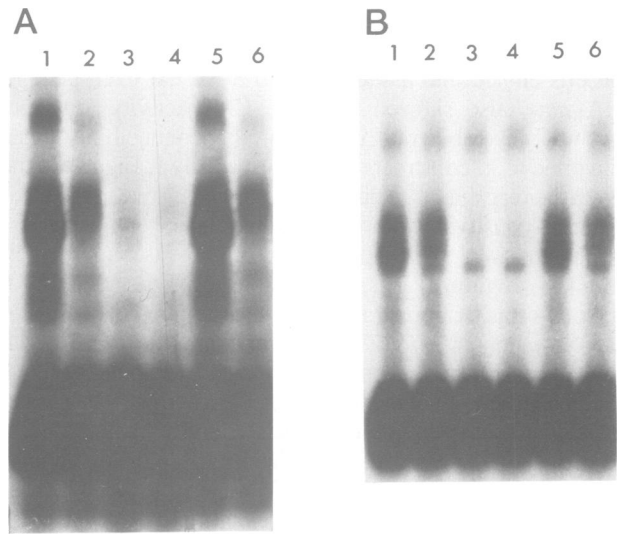


Fig. 6. Gel retardation assays. A synthetic dsDNA fragment corresponding to the 39 bp repeat unit (−127 to −89) was end-labelled, incubated with extracts from untreated (even numbered lanes) or IFN-treated (odd numbered lanes) HeLa cells and analysed by agarose gel electrophoresis (Materials and methods). The whole cell extracts were prepared in 1.2 M (**A**) or 0.4 M salt (**B**) according to Levy *et al.* (personal communication) and Zimarino and Wu (1987) respectively. (**A**) Specific IFN-inducible binding of the 39 bp repeat unit to protein(s). Incubations were carried out with no competitor (**lanes 1** and **2**) or with a 50-fold molar excess of unlabelled 39 bp repeat unit (**lanes 3** and **4**) or non-specific oligonucleotide (**lanes 5** and **6**). The non-specific competitor was the double-stranded synthetic oligonucleotide CCTGGGCGTCTCTGGGAAGTACCG with the single-stranded sequence GATC at each 5′ end. (**B**) Competition between the repeat unit and a 14 bp region for IFN-responsive binding to protein(s). Incubations were carried out with no competitor (**lanes 1** and **2**) or with a 130-fold molar excess of wild-type 14-mer (**lanes 3** and **4**) or mutant 14-mer (**lanes 5** and **6**). The wild-type and mutant 14-mers were synthetic dsDNA fragments corresponding to the sequences GGGAAAATGAAACT and GGGAAAATGACACT respectively. Both double-stranded 14-mers had a single-stranded sequence GATC at each 5′ end.

transfected with these plasmids were tested for the inducibility of 6-16 mRNA (Figure 4). Deletions to positions −414, −318 and −216 did not affect inducibility markedly, while a deletion extending beyond the TATA box to position −7 completely abolished all expression (Figure 4A). A more detailed series of deletions covering the region −199 to −62, extending through the tandem duplication (−89 to −168), provided evidence for a role for each repeat unit in induction (Figure 4B). Inducibility was appreciable (8- to 40-fold) for constructs with deletions 5′ of the first repeat (−199, −196, −171) and extending partially into it (−154). The first of two major losses of inducibility was observed with a deletion to −122. This removed all of the remaining 27 nucleotides (−154 to −127) of the first repeat (including the sequence GGGAAAAT, see below) and five nucleotides of the second, and caused inducibility to be reduced to ~3-fold. Deletions removing another eight (to −114) or 10 (to −112) nucleotides had no further effect, but the removal of an additional eight nucleotides GGGAAAAT (−112 to −104) caused a second and complete loss of inducibility implicating at least a part of this octamer in the induction mechanism. Although uninducible, constructs with further deletions through the 3′ end of the second repeat (to −99 and −88) still showed a basal level of expression. Moreover, this level was higher than that observed (on a longer exposure of the autoradiograph) for constructs with deletions extending to positions (−82 and −74) between the 3′ end of the second repeat and the CCAAT box. The remaining construct with a deletion (to −62) extending through the CCAAT box (at position −63 to −67, opposite strand), showed no detectable expression. The changes in the degree of IFN-inducibility observed in these experiments were reproducible. However, different populations of cells transfected with the same series of plasmids showed variable absolute levels of expression. For example, the apparent increase in the constitutive level of expression observed with the progressive 5′ deletions in Figure 4 was not observed with a separate set of transfectants.

Deletion of the 39 bp *Pst*I fragment (−134 to −95;▽, Figure 2) from the construct shown in Figure 1Aa produced a 6-16 gene lacking the second repeat unit (Materials and methods). Inducibility in L*tk*$^-$ cells transfected with this construct was slightly but reproducibly impaired compared with that observed with the parent plasmid (Figure 4C).

A 101 bp fragment including the tandem duplication can confer inducibility

Taken together, the above results suggested that the tandem duplication can confer inducibility on a reporter gene in a manner relatively independent of position and orientation. As an initial test of this, a construct carrying the HSV*tk* gene, including its TATA and CCAAT boxes and Sp1 binding

Table I. Homologies in 5′-flanking regions of transcriptionally regulated genes

Source	Inducer	Sequence		Position
Human 6-16 (this paper)	IFN	G G G A A A A T G A A A C T	*	−112 to −99
				−153 to −140
Human ISG-54[a]	IFN	G G G A A A G T G A A A C T	*	−87 to −100
Human ISG-15 (Reich *et al.*, 1987)	IFN	G G G A A A C C G A A A C T	*	−108 to −95
Human IP-10 (Luster and Ravetch, 1987)	IFN	T G G A A A G T G A A A C C		−222 to −209
Human Factor B (Wu *et al.*, 1987)	IFN	A G G A A A C A G A A A C T		−127 to −140
Human 2-5A synthetase[b]	IFN	A G G A A A G T G C A A A G		−131 to −118[c]
		A G G A A A C G A A A C C A		−100 to −87[c]
Mouse 202 (Samanta *et al.*, 1986)	IFN	G G G A A A T T G A A A G C		−144 to −131
Human β-IFN (Goodbourn *et al.*, 1985, 1986)	dsRNA	A G G A A A A C T G A A A G		−92 to −79
		G G G A A A T T C C T C T G	*	−64 to −51
Human α-IFN (Ryals *et al.*, 1985)	dsRNA	A G G A A A G C A A A A A C	*	−99 to −86
		T G G A A A G T G G C C C A	*	−80 to −67
Human c-fos (Treisman, 1985)	Serum	T G G A A A C C T G C T G A		−282 to −295
Human HSP gene (Wu *et al.*, 1986)	Serum	G G G A A A A G G C G G G T	*	−55 to −42
SV40 enhancer core (Weiher *et al.*, 1983)		T G G A A A G T C C C C A G	*	−174 to −161
				−246 to −233
HIV/LTR (Nabel and Baltimore, 1987)	PMA/PHA	T G G A A A G T C C C C A G	*	−79 to −92
		C G G A A A G A C C C T T G	*	−93 to −106
CMV enhancer (Boshart *et al.*, 1985)		T G G A A A G T C C C G T T		−93 to −106
		T G G A A A T C C C C G T G		−156 to −169
		A G G A A A G T C C C A T A		−261 to −274
		T G G A A A G T C C C T A T		−412 to −425

The sequences shown are from 5′-flanking DNA of selected human, murine or viral genes known to be regulated at the level of transcription. The positions of these sequences relative to the transcriptional initiation sites are indicated. Analyses of 5′-flanking DNA have been carried out demonstrating that these sequences fall in regions of functional significance. More detailed analyses of deletions and point mutations have directly established the importance of the starred sequences in regulation. Motifs discussed in the text are boxed or underlined. ISG, interferon stimulated gene; HSP, heat shock protein; HIV, human immunodeficiency virus; LTR, long terminal repeat; PMA, phorbol myristate acetate; PHA, phytohaemagglutinin; CMV, human cytomegalovirus.

[a]D.Levy, N.Reich and J.Darnell, Rockefeller University, USA, personal communication.

[b]J.Chebath, Weizmann Institute of Science, Israel, and M.Rutherford and B.Williams, University of Toronto, Canada, personal communication.

[c]Position numbers relative to the A of the initiation codon.

sites, was linked to two head to head copies of a 101 bp fragment which includes the tandem duplication, as shown in Figure 5A. Northern analysis of RNA from mouse cells stably transfected with this construct showed the *tk* transfect to be IFN-inducible (Figure 5B).

IFN-modulated binding of protein(s) to the repeat unit and competition with a subregion oligonucleotide

A synthetic dsDNA fragment corresponding to the 39 bp repeat unit (−127 to −89) was radioactively labelled and tested for its ability to bind proteins as judged by a gel retardation assay. Bands of reduced electrophoretic mobility (compared with the free fragment) were detected with extracts of both untreated and IFN-treated HeLa cells (Figure 6A). Some of these bands were of greater intensity when extracts from IFN-treated cells were used, indicating IFN-modulated interactions between the repeat unit and at least one protein. These interactions were specific since the presence of excess unlabelled repeat unit prevented the formation of detectable IFN-induced bands, whereas excess of an unrelated sequence did not (Figure 6A, lanes 3−6).

Based on sequence comparisons and deletion analyses of 6-16 and other IFN-inducible genes (see Table I and Discussion), a 14 bp subregion of the repeat unit (GGGAAA-ATGAAACT) was predicted to be of particular importance in the IFN response of 6-16. Accordingly, a double-stranded DNA fragment corresponding to this subregion was synthesized and shown to compete with the labelled repeat unit for IFN-modulated binding in the gel retardation assay (Figure 6B, lanes 1−4). A similar DNA fragment containing a single modification (GGGAAAATGACACT) did not compete under identical conditions (Figure 6B, lanes 5 and 6).

Discussion

The analyses presented here have demonstrated the importance of sequences within a tandem duplication (−89 to −168, Figure 2) in mediating the IFN response of the 6-16 gene (Figures 3−6). The tandem duplication (two copies head to head, single copy not tested) conferred inducibility on the HSV*tk* gene (Figure 5). Similarly, a variety of DNA fragments containing the tandem duplication conferred inducibility on the bacterial CAT gene in a manner qualitatively independent of orientation and distance (Figure 3). Only one copy of the repeat unit of the tandem duplication was required to retain inducibility, although both may be required for full induction (deletion analysis, Figure 4). Similarly, in a gel retardation assay, a single copy of the repeat unit is sufficient to interact with protein(s) in an IFN-responsive manner (Figure 6A).

The deletion analysis (Figure 4) suggests further that the tandem duplication acts on a promoter region itself capable of low level constitutive expression in L cells. Two reductions in such expression, apparent on deletion through this region, define sequences (−88 to −82 and −71 to −62) which must each contain the 5′ limit of a promoter element. The −71 to −62 region includes the inverted CCAAT box which appears to be absolutely required for transcription

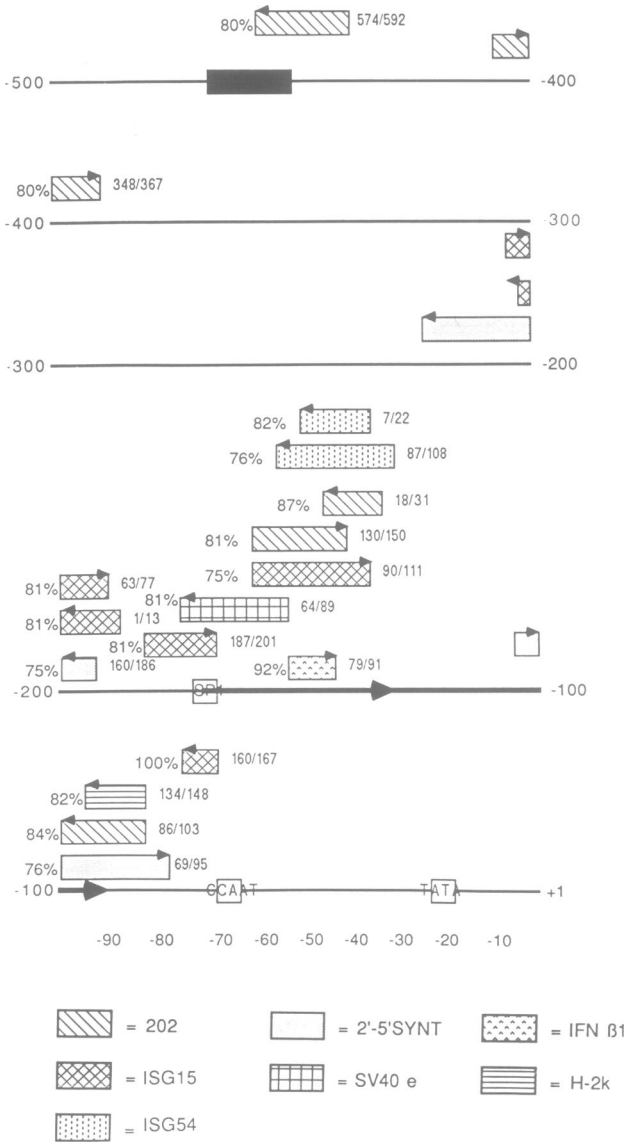


Fig. 7. Homologies in 5′-flanking sequences between 6-16 and other eukaryotic genes. The −500 to +1 region of 6-16 is shown as lines with a −449 to −465 protein binding region (black box; Y.Chernajovsky and G.R.Stark, unpublished results) and the TATA box, CCAAT box and potential Sp1 binding site (open boxes) indicated. Regions of homology with other genes are represented as shaded boxes. The arrows on top of these boxes indicate homology with the plus (→) or minus (←) strands. The percentage on the left of each box indicates the degree of homology; the numbers on the right indicate the position of this homology relative to the transcriptional initiation site for the homologous gene (except for the 2′-5′SYNT where position numbers are relative to the A residue of the initiation codon). Homology to the following genes is shown: 202, an IFN-inducible mouse gene of unknown function (Samanta *et al.*, 1986); ISG-15 (Reich *et al.*, 1987) and ISG-54 (Levy, Reich and Darnell, personal communication), two IFN-inducible human genes of unknown function: 2′-5′SYNT, the IFN-inducible gene for a 2′-5′-A oligoadenylate synthetase (Benech *et al.*, 1985; M.Revel and B.Williams, personal communication); SV40e, the enhancer upstream of the gene for the SV40 T antigen (Reddy *et al.*, 1978); the IFN β1 gene (Goodbourn *et al.*, 1985); the mouse H-2k gene (Kimura *et al.*, 1986).

(Figure 4 and Y.Chernajovsky and G.R.Stark, unpublished results).

Transcriptional control is therefore a major element in induction of the 6-16 gene by IFN and is mediated largely by sequences in the 5′-flanking region of the gene. A general comparison of these sequences with other known sequences (Figure 7) reveals, in addition to the CCAAT box and a potential Sp1 binding site (Dynan *et al.*, 1985), homologies with the upstream regions of a number of IFN-inducible genes and with the enhancers for SV40 and the human β-IFN gene. The majority of these homologies involve sequences within or adjacent to the tandem duplication. One intriguing exception is the homology with the mouse 202 gene (Samanta *et al.*, 1986), involving a region of 6-16 (−449 to −465) identified as the binding site for a nuclear factor (Y.Chernajovsky and G.R.Stark, unpublished results). The significance of this and of the other homologies involving sequences outside the tandem duplication is not yet clear.

A more detailed analysis of homologies within the tandem duplication is presented in Table I. A sequence of 14 nucleotides shows a high degree of homology with sequences in the 5′-flanking regions of six other IFN-inducible genes and the gene for human β-IFN. Consistent with the significance of these 14 nucleotides is the loss of induction observed on deletion of GGGAAAAT (−112 to −104) from the second repeat which suggests that this octanucleotide must at least define the 5′ limit of a region required for inducibility (Figure 4, lanes 21−24). Confirmation of the importance of the 14 nucleotide sequence came from the gel retardation assays in which it was shown to compete with the 39 bp repeat unit for formation of IFN-modulated complexes (Figure 6B). Moreover, analyses of the 5′-flanking sequence of the Type I IFN-inducible genes ISG-15 and ISG-54 have highlighted sequences (Table I) homologous to the 14-mer of 6-16 as being of particular importance in the IFN response (Levy, Reich and Darnell, personal communication; Reich *et al.*, 1987). It is intriguing that the motif GGAAA present in this sequence is found in several well characterized enhancers (Table I, boxed) whereas the more extensive homology reflected by the motif TGAAACT (underlined, Table I), which lies only one residue 3′ of GGAAA, is observed only in the IFN-inducible and β-IFN genes. The significance of the TGAAACT motif is apparent from the loss of competition in the band shift assay on mutating its central A residue (Figure 6B). The TGAAACT motif contains the best homology (GAAACT) with a highly conserved region of the consensus sequence (opposite strand: AGAGA$^{G}/_{A}$GAA ACTGCNGAGGTN$^{G}/_{C}$NGAA) of Friedman and Stark (1985). Interestingly, in that sequence, which is present in the 5′ regions of genes induced less well than 6-16 by α/β-IFN, the GAAACT motif is not associated with a GGAAA motif. A different motif (GTCCCC) is conserved directly 3′ of the GGAAA in the viral enhancers (broken underline, Table I) and, in the case of HIV, this motif is known to be functionally significant (Nabel and Baltimore, 1987).

On the basis of these observations one can postulate a model in which a common motif (e.g. GGAAA) and a more specific motif (e.g. TGAAACT in 6-16) together form the recognition site for a particular, inducer-responsive, positively acting protein. This might be one of several evolutionarily related proteins responding to different inducers, depending on the nature of the second motif, or to the same inducer but with different kinetics or degrees of responsiveness depending the nature of the first motif. Other models can be devised and a full understanding of how these sequences function will have to await the results of a more detailed analysis. Nevertheless, it is reasonable to conclude

that the combined motifs define, or form part of, a response element within the tandem duplication of 6-16. This element has only partial homology to the consensus sequence previously described in most detail for the class I MHC genes (Friedman and Stark, 1985; Israel *et al.*, 1986, 1987; Korber *et al.*, 1987).

There is some variation between the IFN-inducible genes in the exact sequences of the combined motifs (Table I) and substantial variations in the degree of their inducibility by Type I IFNs. Thus the Factor B and IP-10 genes are only poorly induced by these IFNs, and the motifs as they occur in the β-IFN gene (Table I) do not confer IFN inducibility. Similarly, the sequence GAAATAGAAACT (−88 to −77) immediately 3′ of the tandem duplication of 6-16, which has strong homology to the combined motifs (Table I), does not confer inducibility on the 6-16 gene in the absence of the tandem duplication (Figure 4, lanes 23−28). It is possible that the context of the combined motifs has a profound effect on their function or that small variations in the sequences can modulate the response or even abolish it, as is the case for protein binding for the example of the single base change shown in Figure 6B. While everything is consistent with a major role for the 14 nucleotide sequence, flanking sequences and additional factors may well play an important role in modulating the response. Further analyses of the functional significance of the sequences within and surrounding the tandem duplication will be of obvious interest. Additional characterization of nuclear proteins binding to these regions is in progress.

Materials and methods

Growth of cells and IFN treatment

All cells were grown as monolayers in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) heat-inactivated fetal calf serum and, in the cases of stably transfected cells, the selective reagents hypoxanthine, aminopterin and thymidine (i.e. HAT medium; Littlefield, 1964) or hygromycin (400 μg/ml). Mouse cells were treated with recombinant IFN A/D (*Bgl*), a hybrid of two human α-IFNs active on both human and mouse cells (2×10^8 IU/mg protein, Rehberg *et al.*, 1982), supplied by Dr Sidney Pestka, Roche Institute of Molecular Biology, Nutley, NJ, USA. Human cells were treated with Wellferon, a highly purified mixture of human α-IFNs (10^8IU/mg protein, Allen *et al.*, 1982) supplied by Dr K.Fantes, Wellcome Research Laboratories. IFN treatment was at 300 IU/ml. Stable transfectants to be subjected to Northern analysis were grown to ~50% confluence and, unless stated otherwise, for a further 22−24 h in the presence (or continued absence) of IFN. Stable transfectants to be assayed for CAT activity were seeded at 4×10^4 cells per 16-mm diameter well and incubated for 60 h, the last 12, 24 or 48 h of which were in the presence (or continued absence) of IFN.

Recombinant DNA techniques

Standard techniques for the treatment of DNA with restriction endonucleases, T4 DNA ligase, nuclease *Bal*31 and the large fragment of DNA polymerase I (Klenow fragment) were used (Maniatis *et al.*, 1982). Plasmid DNA was prepared by the alkaline lysis procedure (Birnboim and Doly, 1979) followed by two fractionations by ultracentrifugation through CsCl. Fragments of DNA were purified from agarose gel slices using glass beads (Porter *et al.*, 1983).

Transfection procedures

Cells were seeded at 10^6 per 90-mm diameter dish and incubated overnight. DNA (20 μg) was added in the form of a calcium phosphate precipitate. After 8−16 h the cells were treated with glycerol (10% v/v) and, in the case of stable transfections, sodium butyrate (Kelly *et al.*, 1986). Cells to be assayed for the transient expression of CAT activity were grown in the presence (or continued absence) of IFN for the 12 h following treatment with glycerol. Unless stated otherwise, stable transfectants of L*tk*$^-$ cells were selected in HAT medium after cotransfection of a plasmid carrying the *tk* gene (pAGO; Colbère-Garapin *et al.*, 1979) and the plasmid of interest at a molar ratio of ~1:10. Colonies of stably transfected cells, routinely obtained at a frequency of $\sim 10^{-4}$, were pooled and maintained as mixed populations. HeLa cells and some L*tk*$^-$ cells were stably transfected to hygromycin-resistance by a similar procedure except that the plasmid pSV2hygro [based on the plasmid pY3 of Blochlinger and Diggelmann (1984) and supplied by Dr K.Blochlinger, Swiss Institute for Experimental Cancer Research], was used instead of pAGO and selection was in hygromycin (400 μg/ml).

Preparation of cytoplasmic RNA from HeLa and L cells

Cells ($\sim 10^7$) were harvested, resuspended at 4°C in 200 μl lysis buffer (150 mM NaCl; 10 mM Tris−HCl, pH 7.9; 1.5 mM $MgCl_2$; 0.65% (v/v) Nonidet P-40) and nuclei pelleted by centrifugation. Supernatants were made to 0.5% (w/v) SDS and 1 mM EDTA and extracted, at room temperature, twice with 200 μl phenol/chloroform (1:1 by volume) and once with 200 μl chloroform. RNA in the aqueous phase was ethanol-precipitated, washed (70% ethanol) and resuspended in 50−100 μl TE (10 mM Tris−HCl, pH 7.2; 1 mM EDTA).

Northern analyses

Aliquots (20 μg) of cytoplasmic RNA were glyoxalated, fractionated in 1.4% agarose gels and transferred to nylon filters (Pall Corp, McMaster and Carmichael, 1977). These were baked at 80°C for 2 h and prehybridized for 1−24 h at 42°C in hybridization solution [50% (v/v) formamide; 5 × SSPE; 0.25% (w/v) low fat dried milk; 0.1% (w/v) SDS]. Fragments of 6-16 cDNA (Kelly *et al.*, 1986) or HSV*tk* DNA (1.8 kb *Pvu*II fragment) were labelled to $>10^8$ c.p.m./μg by the random priming technique (Feinberg and Vogelstein, 1982) and used to probe for 6-16 and *tk* mRNA respectively. A probe for total cytoplasmic RNA, dominated by cDNA to rRNA and therefore suitable as a probe for rRNA, was prepared as follows; cytoplasmic RNA (1 μg) was boiled for 3 min, chilled on ice and incubated in 50 mM Tris−HCl, pH 8.3; 6 mM $MgCl_2$; 40 mM KCl; 1 mM DTT containing dATP, dGTP and dTTP (50 μM each), [α-^{32}P]dCTP (Amersham; 50 μCi, 6000 Ci/mmol) and reverse transcriptase (Pharmacia; 20 units) at 37°C for 1 h. Following phenol extraction, ethanol precipitation and incubation in 0.1 M NaOH for 1 h at 70°C, the cDNA products were neutralized and ethanol precipitated. Probes were boiled and hybridized to filters in hybridization solution ($>10^6$ c.p.m./ml, for 6-16 and *tk* probes, $>10^5$ c.p.m./ml for the rRNA probe) at 42°C for 12−24 h. Filters were washed twice (15−30 min for each wash) in 2 × SSPE, 0.1% (w/v) SDS (500 ml) at 50°C and blotted dry. Autoradiography and the densitometry of autoradiograms have been described (Kelly *et al.*, 1986). The amount of RNA in each track was normalized by comparing the signals generated by the rRNA probe. Only minor variations were observed and these were taken into account when estimating the degree of 6-16 mRNA inducibility.

Nucleotide sequencing of the 5′ end of the 6-16 gene

DNA from the region of interest (Figure 2) was digested with *Alu*I, *Hae*III or *Sau*3A and cloned into the M13 vector mp18 (Norrander *et al.*, 1983). Random clones were sequenced by the chain termination method (Sanger, 1981). Certain fragments in the region of the tandem duplication were not cloned into M13 at a detectable frequency. Fragments from this region were therefore cloned into pGEM1 (Promega Biotech) and sequenced as dsDNA according to the method of Hattori and Sakaki (1986).

CAT assays

Transiently transfected cells were scraped from each 90-mm dish, pelleted by centrifugation and lysed by resuspension in lysis buffer (100 μl). Stable transfectants were lysed by the addition of lysis buffer (100 μl) to each 16-mm culture well. Nuclei were removed by centrifugation and supernatants were assayed for their ability to catalyse the transfer of [^{14}C]acetyl groups from [^{14}C]acetyl coenzyme A to unlabelled chloramphenicol as described (Sleigh, 1986). Where necessary, lysates were diluted to ensure the assay was in its linear range. Activity is expressed (Figure 3) as % transfer/h/30 μl undiluted lysate.

Gel retardation assays

Extracts were prepared from monolayer cultures of HeLa cells as described by Levy and J.Darnell (personal communication, Figure 6A) or by Zimarino and Wu (1987) (Figure 6B). The cells were either not treated or treated for 2 h with Wellferon (300 IU/ml). Portions (10 μl) of nuclear extract (3 mg/ml protein) were pre-incubated with poly(dI•dC)−poly(dI•dC) (Pharmacia, 1.3 μl of 5 mg/ml) for 10 min at room temperature. Reagents were then added in the form of a single solution (9 μl) to the following final concentrations: yeast tRNA (0.5 mg/ml), pdN_5 (Pharmacia, 0.5 mg/ml), fragmented *Escherichia coli* DNA (0.25 mg/ml), BSA (2 mg/ml), end-

labelled oligonucleotide (10 000 c.p.m., 1 ng) and Ficoll (4%, w/v). The mixture was incubated at room temperature for 20 min and separated on a 1% (w/v) agarose gel in 45 mM Tris/borate buffer, pH 8. The gel was dried on Whatman DE81 paper and autoradiographed.

Construction of plasmids

An 8.7 kb *Xho*I fragment carrying the 6-16 gene was subcloned from the cosmid 10.3 (Kelly *et al.*, 1986) into the *Sal*I site of pGEM2 (Promega Biotech) to form the plasmid shown in Figure 1Aa. This was digested with *Sma*I and *Bgl*II and religated at low concentration to form the plasmid shown in Figure 1Ab.

Nuclease *Bal*31 digests were performed on the plasmid shown in Figure 1Aa after it had been linearized with *Bgl*II. Digests were treated with *Sma*I to remove any remaining human DNA upstream of the *Bgl*II site, and re-circularized. The extent of *Bal*31 digestion downstream of the *Bgl*II site in each of the resulting constructs was determined by sequence analysis of dsDNA (Hattori and Sakaki, 1986) using an oligonucleotide primer homologous to the bacteriophage SP6 promoter.

A 6-16 gene lacking the second of the two repeat units was prepared as follows. The plasmid shown in Figure 1Aa was digested with *Bgl*II and *Kpn*I to generate two fragments (~10.2 and ~1.5 kb). Both of these were purified and the smaller was digested with *Pst*I to generate *Bgl*II/*Pst*I, *Pst*I/*Pst*I and *Pst*I/*Kpn*I fragments of 436, 39 and ~1000 bp respectively. The *Bgl*II/*Pst*I and *Pst*I/*Kpn*I fragments were purified, discarding the *Pst*I/*Pst*I fragment, and ligated to the ~10.2 kb *Bgl*II/*Kpn*I fragment to generate the desired plasmid.

The vectors pSVOcat and $pA_{10}cat_2$ (Gorman *et al.*, 1982; Laimins *et al.*, 1982; Rosenthal *et al.*, 1983) were linearized at their unique cloning sites (*Hin*dIII and *Bgl*II respectively) 5′ of the CAT gene. The ends of the cloning site for pSVOcat were filled by treatment with the Klenow fragment. A 564-bp *Bgl*II/*Bam*HI fragment (Figure 2, −603 to −39) was ligated into $pA_{10}cat_2$ (a and b, Figure 3). The ends of a 645-bp *Bgl*II/*Nci*I fragment (Figure 2, −603 to +42) were filled by treatment with Klenow fragment and ligated into pSVOcat (c and d, 3). A 1047-bp *Bgl*II/*Sma*I fragment, consisting of 1040 bp of 6-16 DNA (Figure 2, −603 to +437) and 7 bp of bacteriophage M13mp19 (Norrander *et al.*, 1983) polylinker DNA, was isolated from a recombinant bacteriophage genome, end-filled with Klenow fragment and ligated into pSVOcat (e and f, Figure 3).

The 6-16/HSV*tk* construct (Figure 5) was made as follows: a 101-bp *Nsp*BII/*Fnu*4HI fragment (−174 to −73) carrying the tandem duplication was ligated to a synthetic *Fnu*4HI/*Xba*I adaptor to generate the *Nsp*BII/ *Xba*I fragment. Two copies of this fragment were ligated into the *Sma*I site of the vector pUC18 (Norrander *et al.*, 1983) as shown (Figure 5). The resulting construct was linearized at its unique *Sph*I site, which was then removed by treatment with Klenow fragment in the absence of nucleotides, and ligated to a 1.8-kb *Pvu*II fragment carrying the HSV*tk* gene (Figure 5).

Acknowledgements

We thank James Darnell, Michel Revel, Bryan Williams and their colleagues for communicating their results to us before publication, Enzo Zimarino and Carl Wu for their invaluable help with gel retardation assays, and Iain Goldsmith for synthesizing the oligonucleotides.

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.

Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.

Blochlinger,K. and Diggelmann,H. (1984) *Mol. Cell. Biol.*, **4**, 2929–2931.

Boshart,M., Weber,F., Jahn,G., Dorsch-Hasler,K., Fleckenstein,B. and Schaffner,W. (1985) *Cell*, **41**, 521–530.

Boss,J.M. and Strominger,J.L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9139–9143.

Chebath,J., Merlin,G., Metz,R., Benech,P. and Revel,M. (1983) *Nucleic Acids Res.*, **10**, 1213–1226.

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

Dynan,W.S., Saffer,J.D., Lee,W.S. and Tjian,R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4915–4919.

Feinberg,A.P. and Vogelstein,B. (1982) *Anal. Biochem.*, **132**, 6–13.

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.

Goodbourn,S., Zinn,K. and Maniatis,T. (1985) *Cell*, **41**, 509–520.

Goodbourn,S., Burnstein,H. and Maniatis,T. (1986) *Cell*, **45**, 601–610.

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.

Graves,B.J., Johnson,P.F. and McKnight,S.L. (1986) *Cell*, **44**, 565–576.

Hattori,M. and Sakaki,Y. (1986) *Anal. Biochem.*, **152**, 232–238.

Israel,A., Kimura,A., Fournier,A., Fellous,M. and Kourilsky,P. (1986) *Nature*, **322**, 743–746.

Israel,A., Kimura.A., Kieran,M., Yano,O., Kanellopoulos,J., Le Bail,O, and Kourilsky,P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2653–2657.

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

Kelly,J.M., Porter,A.C.G., Chernajovsky,Y., Gilbert,C.S., Stark,G.R. and Kerr,I.M. (1986) *EMBO J.*, **5**, 1601–1606.

Kimura,A., Israel,A., Le Bail,O. and Kourilsky,P. (1986) *Cell*, **44**, 261–272.

Korber,B., Hood,L. and Stroynowski,I. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3380–3384.

Kusari,J. and Sen,G.C. (1987) *Mol. Cell. Biol.*, **7**, 528–531.

Laimins,L.A., Khoury,G., Gorman,C., Howard,B. and Gruss,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6453–6457.

Larner,A.C., Jonak,G., Chen,Y-S.E., Korant,B., Knight,E. and Darnell, J.E. Jr (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6733–6736.

Littlefield,J.W. (1964) *Science*, **145**, 709–710.

Luster,A.D. and Ravetch,J.V. (1987) *Mol. Cell. Biol.*, **10**, 3723–3731.

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

Maniatis,T., Goodbourn,S. and Fischer,J.A. (1987) *Science*, **236**, 1237–1245.

McMaster,G.K. and Carmichael,G.G. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4835–4838.

Merlin,G., Chebath,J., Benech,P., Metz,R. and Revel,M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4904–4908.

Nabel,G. and Baltimore,D. (1987) *Nature*, **326**, 711–713

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

Porter,A.C.G., Brusilow,W.S.A. and Simoni,R.D. (1983) *J. Bacteriol.*, **155**, 1271–1278.

Reddy,V.B., Thimmappaya,B., Dhar,R., Subramanian,K.N., Zain,B.S., Pan,J., Ghosh,P.K., Celma,M.L. and Weissman,S.M. (1978) *Science*, **200**, 494–502.

Rehberg,E., Kelder,B., Hoal,E.G. and Pestka,S. (1982) *J. Biol. Chem.*, **257**, 11497–11502.

Reich,N., Evans,B., Levy,D., Fahey,D., Knight,E. Jr and Darnell,J.E. Jr (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6394–6398.

Revel,M. and Chebath,J. (1986) *Trends Biochem. Sci.*, **11**, 166–170.

Rosenthal,N., Kress,M., Gruss,P. and Khoury,G. (1983) *Science*, **222**, 749–755.

Ryals,J., Dierks,P., Ragg,H. and Weissmann,C. (1985) *Cell*, **41**, 497–507.

Samanta,H., Engel,D.A., Chao,H.M., Thakur,A., Garcia-Blanco,M.A. and Lengyel,P. (1986) *J. Biol. Chem.*, **261**, 11849–11858.

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

Sleigh,M.J. (1986) *Anal. Biochem.*, **156**, 251–256.

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

Treisman,R. (1985) *Cell*, **42**, 889–902.

Weiher,H., Konig,M. and Gruss,P. (1983) *Science*, **219**, 626–631.

Wu,B.J., Kingston,R.E. and Morimoto,R.I. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 629–633.

Wu,L-C., Morley,B.J. and Campbell,R.D. (1987) *Cell*, **48**, 331–342.

Zimarino,V. and Wu,C. (1987) *Nature*, **327**, 727–730.

Received on August 21, 1987; revised on October 26, 1987

Note added in proof

The sequence data presented in Figure 2 will appear in the EMBL/GenBank/DDJB Nucleotide Sequence Databases under the accession number Y00828.