Volume 11 Number 3 1983 Nucleic Acids Research

Structure and expression of cloned murine IFN-α genes

Gray D.Shaw*, Werner Boll, Hideharu Taira+, Ned Mantei, Peter Lengyel§ and Charles Weissmann

Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland and §Yale University, Department of Molecular Biophysics and Biochemistry, Box 6666 Yale Station, New Haven, CT 06511, USA

Received 23 December 1982; Accepted 14 January 1983

ABSTRACT

The mouse has an interferon-α (MuIFN-α) gene family containing at least four, and likely more than ten members. A segment of mouse chromosomal DNA and cDNAs encoding murine alpha IFNs have been cloned, and the sequence of two MuIFN-α DNAs determined. No intron was found in the chromosomal gene. The two coding sequences produced biologically active IFN when expressed in monkey cells under the control of an SV40 promoter, and in E.coli under the control of the ampicillinase promoter. MuIFN-α1 had no detectable activity on human cells, while MuIFN-α2 was 20% as active on human as on mouse cells.

INTRODUCTION

The analysis of human interferon (IFN) genes by recombinant DNA technology has shown that there exist not less than 13 closely related genes (and 6 or more pseudogenes) for IFNs of the α type (1-3, K. Henco, unpublished results), one (or perhaps more (4,5)) for β type IFN (6-10) and one for IFN-γ (11-14), the product of immune-induced T cells. Human IFNs of the β and γ type are likely glycosylated (15-18) while most if not all the α type IFNs are not (19). The heterogeneity of the IFNs demonstrated by polyacrylamide gel electrophoresis (19) or HPLC (20) could be due not only to the multiplicity of gene products, but also to size heterogeneity resulting from proteolytic cleavage (21) (physiological or during purification) and, in the case of IFN-β and IFN-γ, heterogeneity in regard to glycosylation.

Mouse IFN is also heterogeneous. Purification of IFN from Ehrlich ascites tumor cells infected with Newcastle disease virus yielded three size classes of molecules. A, with about 35-40,000, B, with about 29-35,000 and C, with 20,000 daltons. Amino termi-

nal sequence analysis (22) showed some homology of A and B with human IFN-β (8 of 24 residues), while C had 11 of 20 residues in common with human IFN-α2 (23). Mouse IFN-α could be completely neutralized by a fraction of an antibody preparation directed against human IFN-α (from leukocytes) and vice versa (24), revealing homologies at the protein level. Mouse IFNs of both the α and the β type are glycosylated (25). A recent report describes two cDNA clones, which are thought to represent two distinct murine IFN-β genes (26). Definitive proof for this identification is still outstanding.

We show here that the mouse has an IFN-α gene family comprising not less than 4, but likely more than 10 members, and present the nucleotide sequence of a chromosomal IFN-α gene and of an IFN-α mRNA, as deduced from its cloned cDNA sequence. Both the chromosomal gene, which contains no introns, and the cDNA were expressed in monkey cells, under the control of an SV40 promoter, and in E.coli, under the control of the β-lactamase promoter, to yield biologically active IFNs.

MATERIALS AND METHODS

DNA blotting

Mouse liver DNA was digested with EcoRI. Completeness of digestion was tested by mixing an aliquot of the reaction mixture, containing 1 µg mouse DNA, with 1 µg phage λ DNA and ascertaining that the latter was completely digested. After electrophoresis through a 0.5% agarose gel, DNA was transferred (27,28) to a Sartorius membrane. The filter strips were hybridized (28) at 28°C or 42°C with nick-translated plasmids containing transcribed regions of human IFN-α1 (Hif-2h, refs. 29,30), α2 (SN206, ref. 23) and α4 (KH4.2, K. Henco, unpublished results, cf. ref. 1). The strips were washed in 0.1x to 1x SSC (1x SSC is 0.15 M NaCl, 0.015 M sodium citrate (pH 7)), 0.1% SDS, at 44°C or 67°C , and autoradiographed.

Isolation of a chromosomal IFN-α sequence.

A library of partially MboI-cleaved BALB C mouse embryo DNA in phage λ Charon 28, provided by P. Leder, was screened by the method of Benton and Davis (31), hybridizing (28) at 28°C with the mixture of probes described above and washing in 0.1 x SSC,

0.1% SDS, at 44°C. The single positive clone found was recloned and rescreened twice.

Isolation of IFN-α cDNA clones.

Poly(A)$^+$ RNA from Newcastle disease virus-infected Ehrlich ascites tumor cells was fractionated by sucrose density gradient centrifugation, and IFN mRNA activity determined by the Xenopus oocytes injection technique (32,33). Double-stranded cDNA was made from RNA fractions with the highest activity and cloned in pBR322 (34). Clones with IFN-α related sequences were identified by colony hybridization (35) with a 690 bp HindIII-EcoRI fragment (see Fig. 2) of the mouse chromosomal IFN-α1 DNA as probe.

Expression of mouse IFN-α genes in mammalian cells.

Plasmids ZpBR322(EcoRI)/chrMuIFN-α1/pGS1 and ZpBR322(Pst)/MuIFN-α2-17 (see Results section) were cleaved with EcoRI and HindIII, or with PstI, respectively. After filling in the ends with Klenow DNA polymerase I (9)(P-L Biochemicals), ligation to BamHI linkers (Collaborative Research) (2), and digestion with BamHI, the 690 bp fragment containing the IFN-α1 sequence and the 960 bp fragment containing the IFN-α2 sequence were isolated and ligated to BamHI-cleaved pKCR (36). After cloning in E.coli HB101, plasmids with the desired orientation were identified by restriction mapping. Plasmid DNA was introduced into the SV40-transformed simian cell line COS 1 (37) as described (38).

Expression of mouse IFN-α genes in E.coli.

The plasmid M51, containing a β-lactamase promoter and initator AUG linked to the coding sequence for mature human IFN-α5 (M. Mishina, unpublished, cf. ref. 1), was linearized by partial digestion with DdeI in the presence of ethidium bromide (39), then cleaved with BamHI. A 590 bp BamHI-DdeI fragment, containing the β-lactamase promoter and the human IFN sequence up to the codon for amino acid 4, was isolated by agarose gel electrophoresis. A 577 bp DdeI-EcoRI fragment from plasmid ZpBR327 (HindIII)/chrMuIFN-α1/pGS3 coding for amino acids 5-166 of the mouse IFN-α1 gene, was similarly prepared. Ligation of these fragments to a 3.2 kb BamHI-PstI fragment of pBR322 DNA (in which the PstI had been converted to an EcoRI end) yielded pMIFN-α1. As the 4 amino terminal amino acids of human IFN-α5 are identical to those of mouse IFN-α1, the complete sequence of the ma-

ture mouse IFN was restored. Mouse IFN-α2 DNA encoding amino acids 2 to 167 was similarly joined to a mature, human IFN-α2 segment linked to the β-lactamase promoter, at the Sau3a site between the codon for amino acids 1 and 2 which is common to both IFNs.

Interferon assay and antibody neutralization.

IFN was assayed on mouse L_{929} and human Hep2 cells by a cytopathic effect reduction assay (40), challenging with Mengo virus. IFN activity on Hep2 cells is in international units. Activity on L_{929} cells is relative to a standard mouse IFN preparation (41) from I. Gresser; one unit equals 4 NIH reference units.

For antibody neutralization, IFN at 50 U/ml was incubated 1 h at 37°C with various dilutions of anti-human IFN serum (42; from K. Cantell) or anti-mouse IFN serum (43; from I. Gresser), and assayed for IFN activity immediately thereafter. The highest dilution neutralizing the IFN activity is taken as the titer of the serum.

Other methods.

DNA sequencing was performed according to Maxam and Gilbert (44). 5'-ends were labeled with γ-^{32}P-ATP and polynucleotide kinase (30), 3'-ends by "filling in" with Klenow DNA polymerase I (9).

RESULTS

1) Mouse DNA has multiple sequences hybridizing to human IFN-α cDNA.

EcoRI-digested mouse DNA was electrophoresed through a 0.5% agarose gel and transferred to a nitrocellulose membrane (27,28). The blots were hybridized to a mixture of ^{32}P-labeled human IFN-α1, IFN-α2 and IFN-α4b DNA fragments (cf. ref. 1) at low stringency and washed under conditions of different stringency. Two mouse DNA samples were run under each condition, one of them containing a linearized HuIFN-α1 cDNA plasmid as internal control. The most stringent washing conditions (0.1 x SSC at 67°C) still allowed the visualization of the human IFN-α DNA, but little hybridization was found to mouse DNA (Fig. 1). Under a variety of non-stringent conditions (0.1 x to 1 x SSC, 44°C) 11 bands of

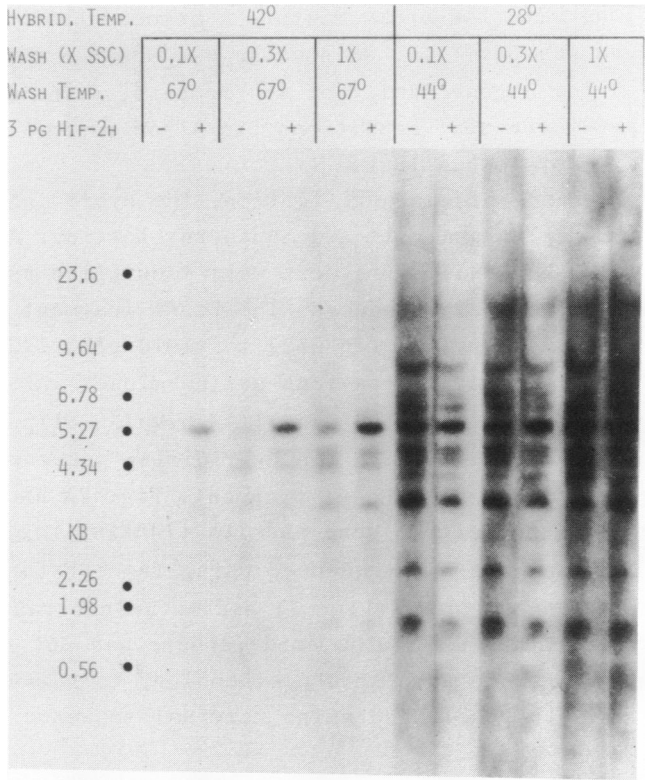


Figure 1. Southern blot analysis of sequences in mouse DNA related to human IFN-α. Alternate samples of EcoRI-digested mouse DNA (10 μg) contained 3 pg linearized human IFN-α1 cDNA plasmid Hif-2H (5.27 kb). Hybridization temperatures and washing conditions are given in the figure. A mixture of human IFN-α DNA fragments was used as probe, as described in the methods section.

differing intensities, ranging from < 2 to 15 kb were revealed in mouse DNA. As the three murine IFN-α coding regions subsequently examined contained no EcoRI site and the chromosomal MuIFN-α gene had no intron, we tentatively conclude that the mouse has 10 or more distinct IFN-α genes.

2) Isolation of a chromosomal mouse IFN-α gene from a genomic library.

A genomic mouse embryo DNA library (provided by P. Leder; average fragment size, 20 kb) was screened (31) under non-strin-

gent conditions with the probe mixture mentioned above; one of 90,000 plaques was positive. This value was about 5 times lower than expected for a gene family of ten members, maybe due to unfavourable screening conditions, to a non-random composition of the library, or to bad luck.

After repeated plaque purification, the hybrid DNA was examined by restriction analysis and Southern blotting. A 1.6-kb EcoRI and a 1.0-kb HindIII fragment were identified as containing the HuIFN-α related sequences. The EcoRI fragment was subcloned into the EcoRI site of pBR322 to yield pBR322(EcoRI)/chr MuIFN-α1/pGS1. The HindIII fragment was subcloned into the HindIII site of pBR327 (pBR327(HindIII)/chrMuIFN-α1/pGS3). The two fragments overlapped, with the hybridizing sequence located in a common 690 bp HindIII-EcoRI fragment. Fig. 2A shows a partial restriction map of the gene and its flanking regions.

A 1230 bp segment was sequenced using the strategy indicated in Fig. 2. The fragment (Fig. 3) had an uninterrupted open reading frame encoding 189 amino acids; there was 76% nucleotide sequence and 62% amino acid sequence homology to human IFN-α1. Comparison with the published amino terminal sequence of mouse

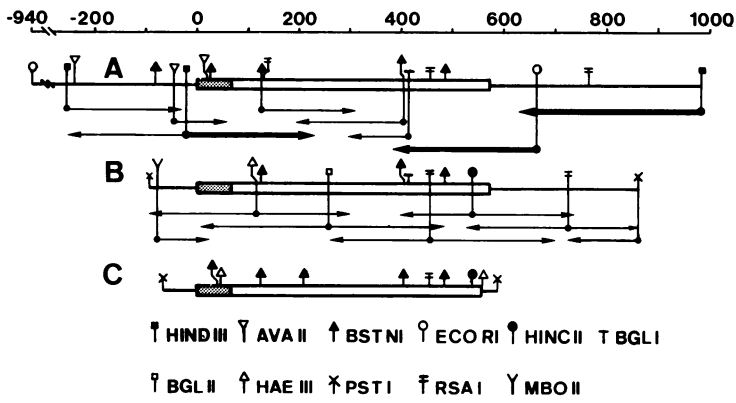


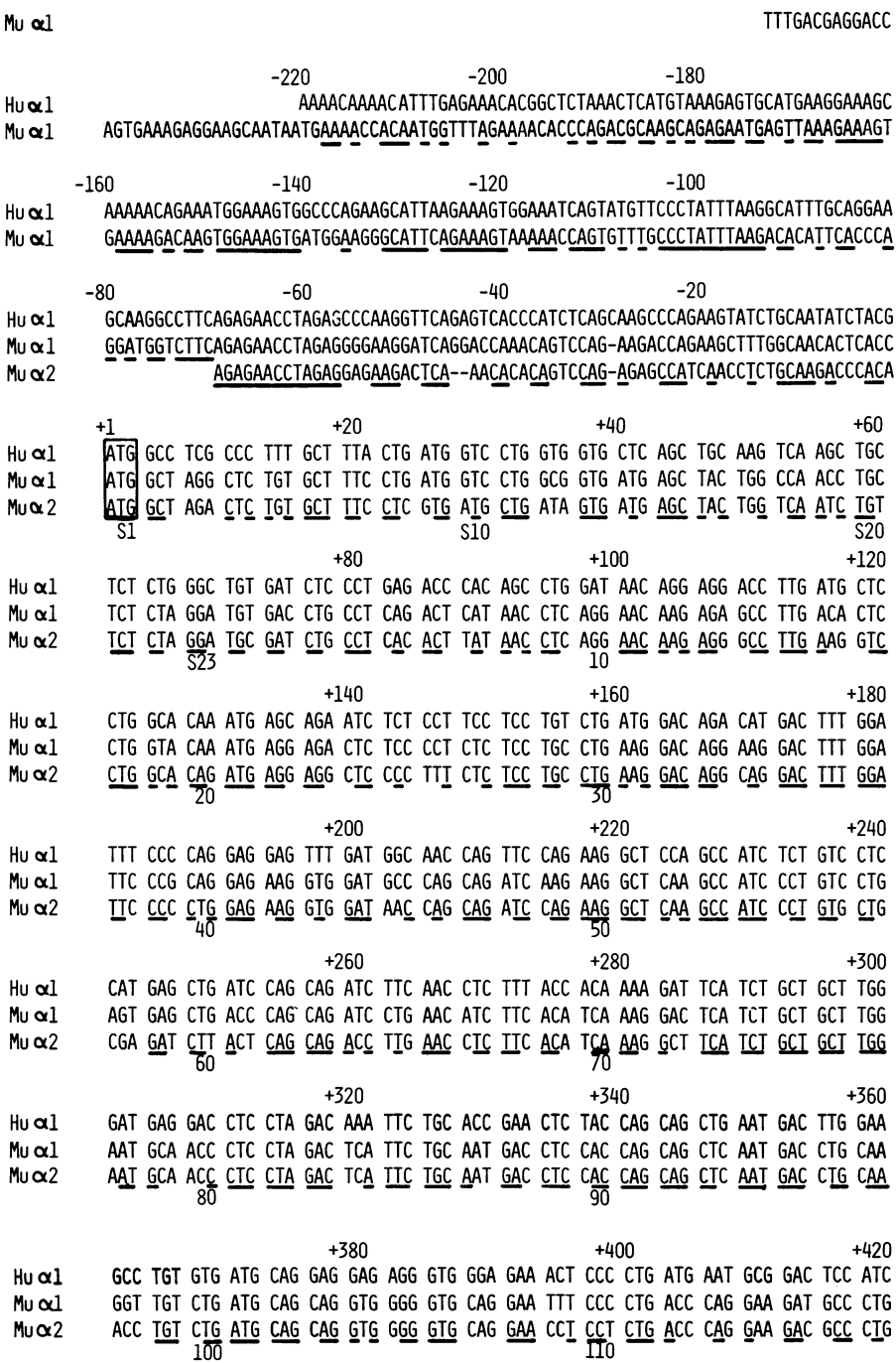
Figure 2. Restriction map and sequencing strategy.
(A) MuIFN-α1, (B) MuIFN-α2, and (C) MuIFN-α3. Filled circles indicate labeled ends, arrows the region sequenced. Bold arrows: Both strands sequenced, using fragments labeled at the 3' and 5' termini of the same fragment, respectively. Empty boxes, mature peptide; shaded boxes, signal sequence.

IFN-C (22; Fig. 4) allowed identification of the mature IFN polypeptide; only 13 of 20 positions (65%) were identical, showing that the cloned gene encoded an IFN-α different from that sequenced by Taira et al. (22). Twenty-three codons upstream from the mature sequence there was an ATG triplet which had to encode the initiator methionine of the preIFN gene, because further upstream there were several termination signals in the reading frame of the gene but no other initiator codon. Thus, the gene encodes a preIFN consisting of a signal sequence of 23 amino acids, followed by a mature sequence of 166 residues, as in the case of all human IFN-α genes except for IFN-α2, which has a mature sequence of only 165 codons. About 370 nucleotides were sequenced downstream from the termination triplet, but no typical "polyadenylation signal" of the type AATAAA (45) or ATTAAA (46,47) was found, perhaps because the 3' non-coding sequence is longer than 400 nucleotides, as is the case for some human IFN-α genes (48). About 250 nucleotides upstream of the coding region were also sequenced; 92 nucleotides 5' to the initiator AUG was the sequence ...CCCTATTTAAG... which comprises a variant of the Goldberg-Hogness sequence (underlined) identical to the one found in almost all human IFN-α genes (49-51; Henco et al., in preparation). Twenty-six nucleotides downstream from the last A residue of the Goldberg-Hogness sequence was the sequence ...AGAGAA... which also occurs in human IFN-α1, and to which the 5' terminus of the human IFN-α1 mRNA has been mapped (49). If the murine IFN-α1 mRNA mapped to the same sequence, it would have 67 nucleotides of 5' non-coding sequence.

Isolation of mouse IFN-α cDNA clones.

Since it was clear from Southern blotting that there were many different mouse IFN-α genes, we decided to isolate at least one IFN-α sequence from a cDNA bank, in order to have a representative which was certain to be derived from an expressed gene.

Poly(A)$^+$ RNA from Ehrlich ascites tumor (EAT) cells induced with Newcastle disease virus (NDV) was centrifuged through a sucrose gradient. The fractions showing IFN mRNA activity were used to generate double-stranded cDNA which was cloned in the PstI site of pBR322 (34). Colonies of E.coli HB101 transformed with the cDNA hybrids were screened with the 690 bp HindIII-


Mu α1 TTTGACGAGGACC
-220 -200 -180
Hu α1 AAAACAAAACATTTGAGAAACACGGCTCTAAACTCATGTAAAGAGTGCATGAAGGAAAGC
Mu α1 AGTGAAAGAGGAAGCAATAATGAAAACCACAATGGTTTAGAAAACACCCAGACGCAAGCAGAGAATGAGTTAAAGAAAGT
-160 -140 -120 -100
Hu α1 AAAAACAGAAATGGAAAGTGGCCCAGAAGCATTAAGAAAGTGGAAATCAGTATGTTCCCTATTTAAGGCATTTGCAGGAA
Mu α1 GAAAAGACAAGTGGAAAGTGATGGAAGGGCATTCAGAAAGTAAAAACCAGTGTTTGCCCTATTTAAGACACATTCACCCA
-80 -60 -40 -20
Hu α1 GCAAGGCCTTCAGAGAACCTAGAGCCCAAGGTTCAGAGTCACCCATCTCAGCAAGCCCAGAAGTATCTGCAATATCTACG
Mu α1 GGATGGTCTTCAGAGAACCTAGAGGGAAGGATCAGGACCAAACAGTCCAG-AAGACCAGAAGCTTTGGCAACACTCACC
Mu α2 AGAGAACCTAGAGGAGAAGACTCA--AACACACAGTCCAG-AGAGCCATCAACCTCTGCAAGACCCACA
+1 +20 +40 +60
Hu α1 ATG GCC TCG CCC TTT GCT TTA CTG ATG GTC CTG GTG GTG CTC AGC TGC AAG TCA AGC TGC
Mu α1 ATG GCT AGG CTC TGT GCT TTC CTG ATG GTC CTG GCG GTG ATG AGC TAC TGG CCA ACC TGC
Mu α2 ATG GCT AGA CTC TGT GCT TTC CTC GTG ATG CTG ATA GTG ATG AGC TAC TGG TCA ATC TGT
S1 S10 S20
+80 +100 +120
Hu α1 TCT CTG GGC TGT GAT CTC CCT GAG ACC CAC AGC CTG GAT AAC AGG AGG ACC TTG ATG CTC
Mu α1 TCT CTA GGA TGT GAC CTG CCT CAG ACT CAT AAC CTC AGG AAC AAG AGA GCC TTG ACA CTC
Mu α2 TCT CTA GGA TGC GAT CTG CCT CAC ACT TAT AAC CTC AGG AAC AAG AGG GCC TTG AAG GTC
S23 10
+140 +160 +180
Hu α1 CTG GCA CAA ATG AGC AGA ATC TCT CCT TCC TCC TGT CTG ATG GAC AGA CAT GAC TTT GGA
Mu α1 CTG GTA CAA ATG AGG AGA CTC TCC CCT CTC TCC TGC CTG AAG GAC AGG AAG GAC TTT GGA
Mu α2 CTG GCA CAG ATG AGG AGG CTC CCC TTT CTC TCC TGC CTG AAG GAC AGG CAG GAC TTT GGA
20 30
+200 +220 +240
Hu α1 TTT CCC CAG GAG GAG TTT GAT GGC AAC CAG TTC CAG AAG GCT CCA GCC ATC TCT GTC CTC
Mu α1 TTC CCG CAG GAG AAG GTG GAT GCC CAG CAG ATC AAG AAG GCT CAA GCC ATC CCT GTC CTG
Mu α2 TTC CCC CTG GAG AAG GTG GAT AAC CAG CAG ATC CAG AAG GCT CAA GCC ATC CCT GTG CTG
40 50
+260 +280 +300
Hu α1 CAT GAG CTG ATC CAG CAG ATC TTC AAC CTC TTT ACC ACA AAA GAT TCA TCT GCT GCT TGG
Mu α1 AGT GAG CTG ACC CAG CAG ATC CTG AAC ATC TTC ACA TCA AAG GAC TCA TCT GCT GCT TGG
Mu α2 CGA GAT CTT ACT CAG CAG ACC TTG AAC CTC TTC ACA TCA AAG GCT TCA TCT GCT GCT TGG
60 70
+320 +340 +360
Hu α1 GAT GAG GAC CTC CTA GAC AAA TTC TGC ACC GAA CTC TAC CAG CAG CTG AAT GAC TTG GAA
Mu α1 AAT GCA ACC CTC CTA GAC TCA TTC TGC AAT GAC CTC CAC CAG CAG CTC AAT GAC CTG CAA
Mu α2 AAT GCA ACC CTC CTA GAC TCA TTC TGC AAT GAC CTC CAC CAG CAG CTC AAT GAC CTG CAA
80 90
+380 +400 +420
Hu α1 GCC TGT GTG ATG CAG GAG GAG AGG GTG GGA GAA ACT CCC CTG ATG AAT GCG GAC TCC ATC
Mu α1 GGT TGT CTG ATG CAG CAG GTG GGG GTG CAG GAA TTT CCC CTG ACC CAG GAA GAT GCC CTG
Mu α2 ACC TGT CTG ATG CAG CAG GTG GGG GTG CAG GAA CCT CCT CTG ACC CAG GAA GAC GCC CTG
100 110

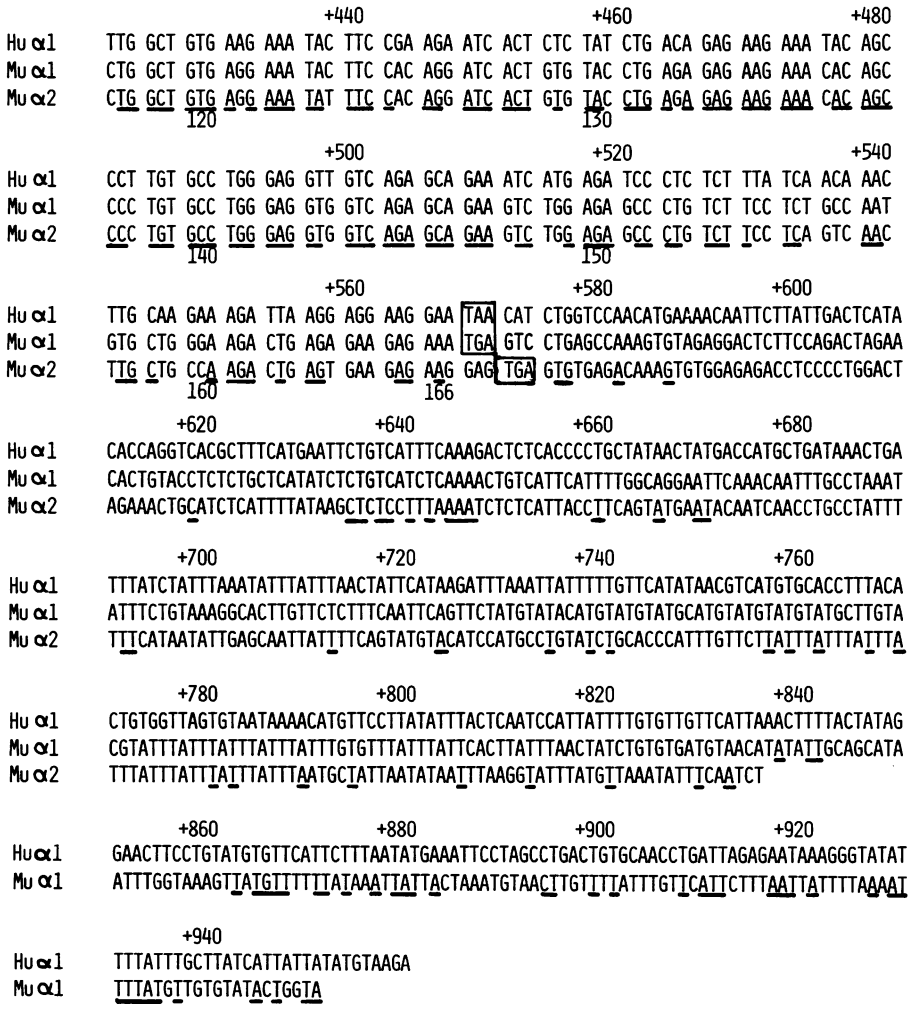


Figure 3. Comparison of human and murine IFN-α nucleotide sequences. The sequence of the human IFN-α1 gene is from (49); murine sequences were determined (44) by the strategy shown in Fig. 2. Initiation and termination codons are boxed. Nucleotides identical in all sequences are underlined.

EcoRI DNA fragment from the mouse chromosomal IFN-α1 clone. The 960-bp insert of one of the positive colonies (ZpBR322(Pst)/ MuIFN-α2-17) was sequenced; it had an open reading frame of 570 nucleotides and was closely related to the MuIFN-α1 described

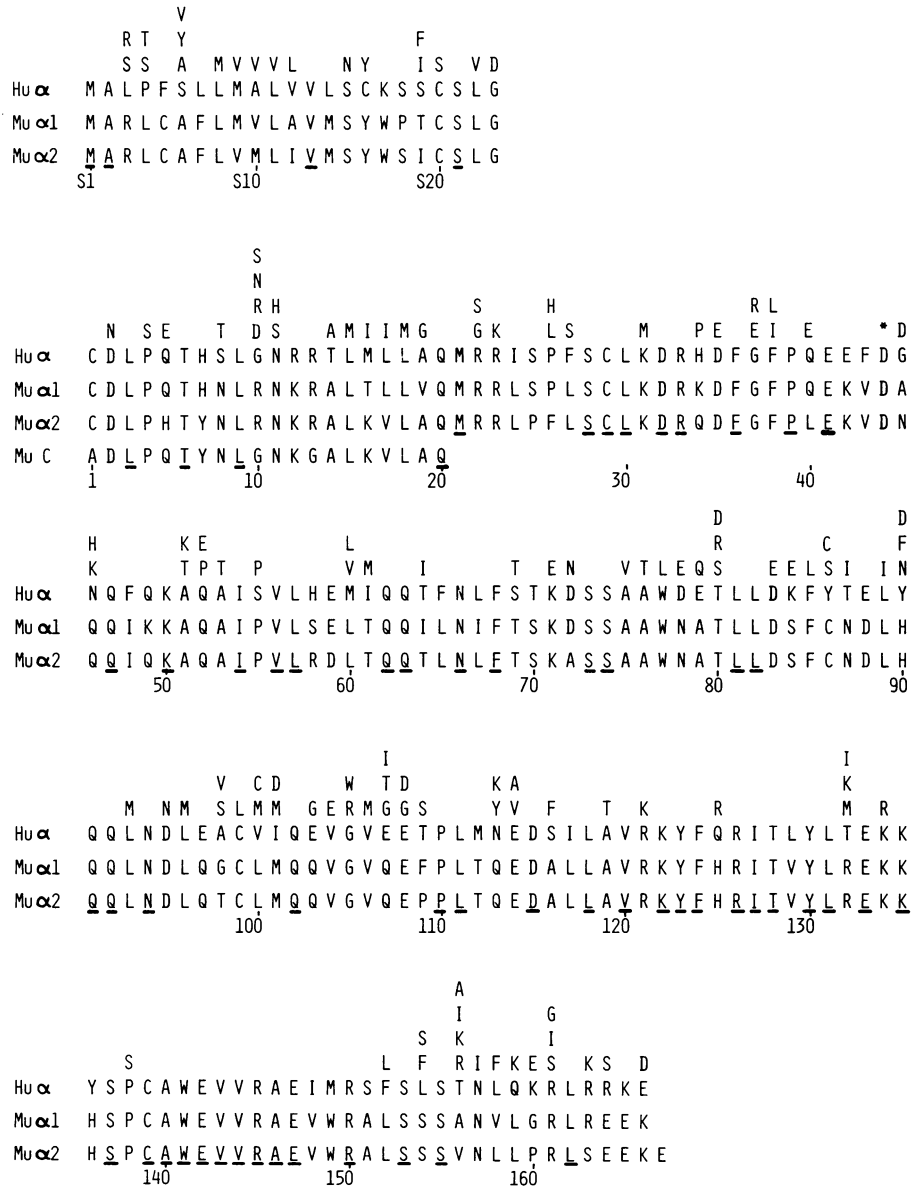


Figure 4. Comparison of IFN-α amino acid sequences. Amino acids are written in one letter notation (61). The human sequence shows all amino acids which occur in each position for the 11 human IFN-genes (3,48,50; K. Henco, unpublished results) listed in Table 2. The sequence of murine IFN-C is from ref. 22. Residues identical in all sequences are underlined.

above (Fig. 3). The sequence for the mature IFN polypeptide MuIFN-α2 had one codon more than MuIFN-α1. Thus, murine preIFN-α2 consisted of a signal sequence of 23 and a mature sequence of 167 amino acids. A further cDNA clone, ZpBR322(Pst)/MuIFN-α3-16 encoded an IFN (MuIFN-α3) different from MuIFN-α1 and α2, as shown by restriction mapping (Fig. 2) and partial sequencing (data not shown).

Expression of the cloned murine interferon coding sequences by transformed monkey cells and E.coli.

To establish that the cloned murine DNA segments in fact encoded biologically active IFNs, they were expressed both in animal cells and E.coli.

The 690 bp HindIII-EcoRI fragment of MuIFN-α1, which contained the entire pre-IFN coding sequence as well as 15 5'-non-coding and 100 3'-non-coding nucleotides, was provided with BamH1 linkers and introduced into the BamH1 site of the eukaryotic expression plasmid pKCR (36), under the control of the SV40 early promoter (see also Haynes and Weissmann, submitted). Monkey COS cells (37) transformed by this plasmid (38) produced up to 13,000 units/ml of antiviral activity (as measured on L929 mouse cells) between the 3rd and the 10th day after transformation, whereafter production declined and ceased. MuIFN-α2 DNA was introduced into pKCR in a like manner and produced similar levels of antiviral activity. Mock transformed cells gave no antiviral activity.

For expression in E.coli, the mature coding sequences of MuIFN-α1 and MuIFN-α2 were joined to the β-lactamase control region of pBR322 such that the initiator AUG of the β-lactamase gene abutted the first codon of the mature IFN sequence. A similar construction carried out with the human IFN-α2 gene led to the production, in E.coli, of mature IFN-α2 (M. Mishina, W. Boll and C.W., quoted in ref. 1). Extracts of E.coli containing these plasmids showed antiviral activity on mouse cells, about 10^3 units/ml culture.

From these experiments it may be concluded that MuIFN-α1 and MuIFN-α2 DNA encode functional murine alpha IFNs.

Some properties of cloned murine alpha interferons.

Murine IFNs from extracts of transformed E.coli or from

transformed monkey cells were assayed on mouse and human cells. Whatever the source, murine IFN-α1 had no detectable effect on human cells (less than 1% the antiviral activity on mouse cells), while murine IFN-α2 (Table 1), either crude or partially purified, was about 20% as active on human cells as on mouse cells. It was reported earlier that different species of human IFN-α can have different target cell specificities (23,52,53). Antiserum against natural murine IFNs completely neutralized the IFNs from cells transformed with cloned MuIFN DNA at a dilution of 1:80 to 1:160, and had no detectable activity against HuIFN-α5. An antiserum against natural human leukocyte IFNs had a titer of 1:100,000 against HuIFN-α5 and 1:600 - 1:1200 against the murine alpha IFNs. The neutralization of natural murine IFN-α by anti-human IFN antisera at a similarly low relative titer has been described by Kawade et al. (24).

Natural expression of MuIFN-α and MuIFN-α2 in Ehrlich ascites tumor (EAT) cells after induction with NDV.

As MuIFN-α2 DNA is a cDNA copy of mRNA, it is clear that

Table 1. The serological cross-activities and relative anti-viral activities on human and mouse cells of MuIFN-α1, MuIFN-α2 and HuIFN-α5.

IFN	Source	Titer on L929 cells	Titer on Hep2 cells	Anti-murine IFN serum titer	Anti-human IFN serum titer
MuIFN-α1	E.coli[a]	1,000	<10	1:160	1:1280
MuIFN-α1	COS cells[b]	1,000	<10	1:80	1:640
MuIFN-α2	E.coli[a]	1,000	300	1:160	1:640
	partially purified[c]	50,000	10,000	N.D.	N.D.
MuIFN-α2	COS cells[b]	1,000	100	1:80	1:640
HuIFN-α5	COS cells[b]	300	10,000	NONE	1:102,400

The serological crossactivities and relative anti-viral activities on human and mouse cells of MuIFN-α1, MuIFN-α2 and HuIFN-α5. (a) IFN activities were measured in extracts of E.coli prepared as described (29) or (b) in tissue culture supernatants. (c) IFN was partially purified from E.coli (s.a. 10^6 units/mg protein) essentially as described in Palva et al. (submitted for publication). Neutralization titers and anti-viral activities were determined as described in the Methods section.

this gene is expressed under natural conditions, at least by NDV-induced EAT cells. In order to determine whether MuIFN-α1 was expressed to a similar extent, poly(A) RNA from NDV-induced EAT cells was analyzed by S_1 mapping (54,55) using specific MuIFN-α1 and MuIFN-α2 probes. The MuIFN-α1 probe was 5' labeled at the EcoRI site, which is not present in MuIFN-α2, and the MuIFN-α2 probe was labeled at the BglII site, which is not present in MuIFN-α1; both probes had known (similar) specific activities. There was a strong signal with the IFN-α2 probe, and a 4-16 times weaker signal with the HuIFN-α1 probe (data not shown), indicating the α1 (or a closely related) gene is expressed only at low level in NDV-induced cells.

DISCUSSION

Two DNAs encoding murine alpha IFNs have been cloned and fully sequenced. MuIFN-α1 was from chromosomal mouse DNA, and MuIFN-α2 from cDNA prepared from mRNA of NDV-induced EAT cells. A further cDNA, MuIFN-α3, was partially sequenced and not further investigated. As in the case of the human alpha and beta IFNs (1,6,9,48,50), no introns could be found in the murine chromosomal gene. Both MuIFN-α1 and MuIFN-α2 expressed antivirally active IFNs in monkey cells and in E.coli. Natural murine IFN-α is glycosylated (25); most likely the murine IFNs made by transformed monkey cells are also glycosylated, while those made by E.coli are not. This view is supported by a comparative analysis of human IFN-γ produced by eukaryotic cells and E.coli, respectively (J. Haynes and C. Weissmann, submitted). As far as they were tested, target cell specificity and serological reactivity were similar whether an IFN was from monkey cells or E.coli.

Comparison of different IFN-α genes.

Fig. 3 shows a comparison of two murine IFN-α sequences and human IFN-α1. About half the residues of the mouse signal sequences differ from those of their human counterpart; only 4 residues are common to all IFN-α sequences. There is more divergence among signal sequences than among the mature IFN sequences, indicating a lesser degree of evolutionary constraint. The mature MuIFN-α1 polypeptide is encoded by 166 codons, as are all

human IFN-α's (except for HuIFN-α2, which has one codon less). Mature MuIFN-α2 is encoded by 167 amino acids, the additional amino acid being the carboxy terminal one, probably due to an insertion of three nucleotides just before the termination triplet. There is a potential glycosylation peptide sequence Asn-Ala-Thr at positions 78-80. HuIFN-β, which is thought to be glycosylated (15-17), but not human alpha IFNs, also have a potential glycosylation sequence in this position.

The coding regions of the two murine IFN-α genes differ in 10.2% of the nucleotide positions, and the cognate mature proteins in 13.8% of the amino acid positions. This divergence in amino acid sequence is similar to that found among human alpha IFNs such as α1 and α2 (23). The 20 first residues of murine IFN-C (22) differ from those encoded by MuIFN-α1 and MuIFN-α2 by 35% and 20%, respectively. They also differ from those of the third murine IFN-α gene we have identified (MuIFN-α3; W. Boll, unpublished results), so that there is structural evidence for at least four distinct MuIFN-α genes.

The murine IFN-α1 and α2 genes were compared to each of the human IFN-α genes, and the divergence (corrected for multiple mutations) of both silent and replacement sites was calculated according to Perler et al. (56) (Table 2). There is considerable scatter in the corrected divergence values when either of the murine IFN sequences is compared with the various human sequences. In particular, the ratio of the values calculated for silent and replacement sites of the same pair fluctuate, between 1.9 and 2.6 for the different mouse-human comparisons, which should not be if the simple "evolutionary clock" model (57) were accurate. In fact, the scatter is likely due to small random variations in divergence which are greatly amplified in the correction calculations of Perler et al. (56). We therefore considered only the average corrected divergence between all murine and human IFN-α genes, namely 55.6% for silent sites and 25.6% for replacement sites. For globin genes, Perler et al. (55) have calculated the time required for the fixation of 1% nucleotide changes between two lines (unit evolutionary period, UEP) as 0.7 million years (MYR) for silent and 10 MYR for replacement sites. Applying these value to IFN-α genes gives divergence times of

Table 2. Divergence of murine and human IFN-α genes.

	Nucleotide Difference (%)		Amino Acid Difference (%)		Corrected Divergence Replacement Sites (%)		Corrected Divergence Silent Sites (%)	
	MuIFN-α1	MuIFN-α2	MuIFN-α1	MuIFN-α2	MuIFN-α1	MuIFN-α2	MuIFN-α1	MuIFN-α2
MuIFN-α1	-	10.2	-	13.8	-	8.4	-	24.6
HuIFN-α1	23.6	26.1	37.6	40.7	22.5	25.4	48.9	55.0
HuIFN-α2	23.8	25.4	39.4	40.4	23.8	25.9	47.0	49.2
HuIFN-α4B	25.4	27.2	40.2	40.7	24.1	25.9	55.8	59.8
HuIFN-α5	24.9	26.5	40.2	41.8	22.6	24.3	55.4	60.1
HuIFN-α6	25.0	26.6	41.3	42.9	24.0	25.6	49.0	56.9
HuIFN-α7	25.7	28.2	39.7	41.3	24.4	27.0	58.2	69.6
HuIFN-α8	23.8	25.8	38.9	39.5	25.5	26.8	49.1	57.8
HuIFN-αF	25.4	27.7	40.7	42.9	24.8	27.1	55.5	63.0
HuIFN-αλ2H	24.9	26.6	40.7	42.3	23.8	27.6	44.4	57.3
HuIFN-αλC1	25.6	27.5	40.7	42.3	24.9	26.8	53.5	57.7
HuIFN-αC	25.6	27.2	39.7	41.3	24.1	26.3	58.1	61.3

The divergence at replacement and silent sites was corrected for multiple mutations by the method of Perler et al. (56). The calculations were performed on an Apple II computer using a program written by John David Weissmann.

about 40 to 260 MYR, respectively. If the divergence of silent sites becomes "saturated" after long divergence times, as has been claimed (56), the higher value should be the more likely one. As the human and mouse lines diverged some 85 MYR ago, the value of 260 MYR is too high. It is quite possible that the scatter in the data are responsible for this apparent inaccuracy of the "evolutionary clock". Nevertheless, two other explanations should be considered: (1) The rate of evolution of some genes may be higher than that of globin genes. For example, the UEP for the replacement sites of growth hormone genes is 4.5 MYR (56), a value closer to that of 3.3 MYR calculated for the IFN-α genes. (2) A large IFN-α gene family may have existed earlier than 250 MYR ago, with some members giving rise to the mouse IFN-α gene family and others to the human family. The fact that the apparent divergence within a species is smaller than that between species (Table 2) could be attributed to gene correction or conversion (58-60) following divergence of the mouse and human lines. This hypothesis leads to the prediction that there could be functional IFN-α genes within one and the same species showing greater divergence than certain IFN-α genes in different species.

The amino acid sequences of 11 human and 2 murine alpha IFNs have 65 completely conserved positions in common, of which 33 are also conserved in HuIFN-β. These conserved positions are clustered between positions 118 to 147 where 22 of 30 (73%) residues are conserved, and 28 to 57 where 13 of 30 (43%) residues are conserved. In contrast, the remainder of the sequence shows only 23% conserved residues.

The two cloned MuIFN-α species show markedly different antiviral activity on human cells, MuIFN-α2 being at least 20 times more active than MuIFN-α1 on human cells, relative to their activity on mouse cells. The two IFNs differ in 23 amino acid positions; among these discordant positions, MuIFN-α2 resembles the human consensus sequence in 4 positions (19,41,49, 67) as indicated in Fig. 4.

Conversely, HuIFN-α1 is 20 times more more active than IFN-α2 on mouse cells (52), and a hybrid HuIFN-α, consisting of amino acids 1 to 63 of IFN-α2 and amino acids 64 to 166 of IFN-α1 shows

an even higher relative activity on mouse cells (52). There are only 5 positions (69,86,101,152,161) in which HuIFN-α1 differs from HuIFN-α2 and agrees with the murine IFN-α sequences between positions 64 and 166.

It will be of interest to determine the minimal number of changes required to modify the target cell specificity of an IFN, as it should help identify the region of the molecule involved in the interaction with the receptor. Sequence comparisons such as the ones outlined above may provide guidance in the preparation of appropriately modified IFNs.

ACKNOWLEDGEMENTS

This work was supported by the Kanton of Zürich, the Schweizerische Nationalfonds, Biogen S.A. and NIH Research Grants No. CA16038 and AI12320. We thank P. Leder for the mouse genomic library, H. Chao for confirming parts of the sequence of MuIFN-α2, J.D. Weissmann for writing a divergence program for the Apple II computer, P. Dierks and J.-I. Fujisawa for γ-^{32}P-ATP, and K. Cantell and I. Gresser for IFN and antibody.

Present address: *Biogen Inc., 241 Binney Street, Cambridge, MA 02142, USA

+Department of Pure and Applied Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153, Japan

REFERENCES

1. Weissmann, C. (1981) In "Interferon 3" (I. Gresser, ed.), Academic Press London, pp. 101-134.
2. Brack, C., Nagata, S., Mantei, N. and Weissmann, C. (1981) Gene 15, 379-394.
3. Weissmann, C., Nagata, S., Boll, W., Fountoulakis, M., Fujisawa, A., Fujisawa, J.-I., Haynes, J., Henco, K., Mantei, N., Ragg, H., Schein, C., Schmid, J., Shaw, G., Streuli, M., Taira, H., Todokoro, K. and Weidle, U. (1982) Phil. Trans. R. Soc. Lond. B299, 7-28.
4. Sagar, A.D., Sehgal, P.B., Slate, D.L. and Ruddle, F.H. (1982) J. Exp. Med. 156, 744-755.
5. Weissenbach, J., Chernajovsky, Y., Zeevi, M., Shulman, L., Soreq, H., Nir, U., Wallach, D., Perricaudet, M., Tiollais, P. and Revel, M. (1980) Proc. Natl. Acad. Sci. USA 77, 7152-7156.
6. Ohno, S. and Taniguchi, T. (1981) Proc. Natl. Acad. Sci. USA 78, 5305-5309.
7. Maeda, S., McCandliss, R., Chiang, T.-R., Costello, L., Levy, W.P., Chang, N.T. and Pestka, S. (1981) In "Developmental Biology: Using Purified Genes" (D.D. Brown, ed.) Academic Press, N.Y., pp. 85-96.
8. Tavernier, J., Derynck, R. and Fiers, W. (1981) Nucl. Acids Res. 9, 461-471.

9. Lawn, R.M., Adelman, J., Franke, A.E. Houck, C.M., Gross, M., Najarian, R. and Goeddel, D.V. (1981) Nucl. Acids Res. 9, 1045-1052.
10. Houghton, M., Jackson, I.J., Porter, A.G., Doel, S.M., Catlin, G.H., Barber, C. and Carey, N.H. (1981) Nucl. Acids Res. 9, 247-266.
11. Gray, P.W., Leung, D.W., Pennica, D., Yelverton, E., Najarian, R., Simonsen, C.C. Derynck, R., Sherwood, P.J., Wallace, D.M., Berger, S.L., Levinson, A.D. and Goeddel, D.V. (1982) Nature 295, 503-508.
12. Derynck, R., Leung, D.W., Gray, P.W. and Goeddel, D.V. (1982) Nucl . Acids Res. 10, 3605-3615.
13. Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G. and Fiers, W. (1982) EMBO J. 1, 953-958.
14. Gray, P.W. and Goeddel, D.V. (1982) Nature 298, 859-863.
15. Knight, E., Jr. (1976) Proc. Natl. Acad. Sci. USA 73, 520-523.
16. Havell, E.A., Yamazaki, S. and Vilček, J. (1977) J. Biol. Chem. 252, 4425-4427.
17. Tan, Y.H., Baraket, F., Berthold, W., Smith-Johannsen, H. and Tan, C. (1979) J. Biol. Chem. 254, 8067-8073.
18. Yip, Y.K., Pang, R.H.L., Urban, C. and Vilček, J. (1981) Proc. Natl. Acad. Sci. USA 78, 1601-1605.
19. Allen, G. and Fantes, K.H. (1980) Nature 287, 408-411.
20. Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R. S., Waldman, A.A. and Pestka, S. (1979) Proc. Natl. Acad. Sci USA 76, 640-644.
21. Levy, W.P., Rubinstein, M., Shively, J., Del Valle, U., Lai, C.-Y., Moschera, J., Brink, L., Gerber, L., Stein, S. and Pestka, S. (1981) Proc. Natl. Acad. Sci. USA 78,6186-6190.
22. Taira, H., Broeze, R., Jayaram, B.M., Lengyel, P., Hunkapiller, M.W. and Hood, L.E. (1980) Science 207, 528-530.
23. Streuli, M., Nagata, S. and Weissmann, C. (1980) Science 209, 1343-1347.
24. Kawade, Y., Watanabe, Y., Yamamoto, Y., Fujisawa, J., Dalton, B.J. and Paucker, K. (1981) Antiviral Res. 1, 167-178.
25. Fujisawa, J., Iwakura, Y. and Kawade, Y. (1978) J. Biol. Chem. 253, 8677-8679.
26. Skup, D., Windass, J.D., Sor, F., George, H., Williams, B.R. G., Fukuhara, H., De Maeyer-Guignard, J. and De Meyer, E. (1982) Nucl. Acids. Res. 10, 3069-3084.
27. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
28. Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-3687.
29. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K. and Weissmann, C. (1980) Nature 284, 316-320.
30. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. and Weissmann, C. (1980) Gene 10, 1-10.
31. Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
32. Reynolds, F.H., Jr., Premkumar, E. and Pitha, P.M. (1975) Proc. Natl. Acad. Sci. USA 72, 4881-4885.
33. Colman, A. and Morser, J. (1979) Cell 17, 517-526.
34. Land, H., Grez, M., Hauser, H. Lindenmaier, W. and Schütz, G. (1981) Nucl. Acids Res. 9, 2251-2265.
35. Hanahan, D. and Meselson, M. (1980) Gene 10, 63-67.
36. O'Hare, K., Benoist, C. and Breathnach, R. (1981) Proc. Natl. Acad. Sci. USA 78, 1527-1531.

37. Gluzman, Y. (1981) Cell 23, 175-182.
38. Frost, E. and Williams, J. (1978) Virol. 91, 39-50.
39. Parker, R.C., Watson, R.M. and Vinograd, J. (1977) Proc. Natl. Acad. Sci. USA 74, 851-855.
40. Stewart, W.E. II (1979) In "The Interferon System", Springer, Berlin.
41. Tovey, M.G., Begon-Lours, J. and Gresser, I. (1974) Proc. Soc. Exp. Biol. Med. 146, 809-815.
42. Mogensen, K.E., Pyhalä, L. and Cantell, K. (1975) Acta Path. Microbiol. Scand. 83B, 443-450.
43. Gresser, I., Tovey, M.G. Bandu, M.T. Maury, C. and Brouty-Boye, E. (1976) J. Exp. Med. 144, 1305-1315.
44. Maxam, A.M. and Gilbert, W. (1980) In "Methods in Enzymology" (K. Grossman and K. Moldave, eds.), Academic Press, N.Y., 65, 499-560.
45. Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211-214.
46. Hagenbüchle, O., Bovey, R. and Young, R.A. (1980) Cell 21, 179-187.
47. Nunberg, J.H., Kaufman, R.J., Chang, A.C.Y., Cohen, S.N. and Schimke, R.T. (1980) Cell 19, 355-364.
48. Goeddel, D.V. Leung, D.W., Dull, T.J., Gross, M., Lawn, R. M., McCandliss, R., Seeburg, P.H., Ullrich, A., Yelverton, E. and Gray, P.W. (1981) Nature 290, 20-26.
49. Nagata, S., Mantei, N. and Weissmann, C. (1980) Nature 287, 401-408.
50. Lawn, R.M., Adelman, J., Dull, T.J., Gross, M., Goeddel, D. V. and Ullrich, A. (1981). Science 212, 1159-1162.
51. Lawn, R.M., Gross, M., Houck, C.H., Franke, A.E., Gray, P.W. and Goeddel, D.V. (1981) Proc. Natl. Acad. Sci. USA 78, 5435-5439.
52. Streuli, M., Hall, A., Boll, W., Stewart, W.E. II, Nagata, S. and Weissmann, C. (1981) Proc. Natl. Acad. Sci. USA 78, 2848-2852.
53. Weck, D.K., Apperson, S., Stebbing, N., Gray, P.W., Leung, D., Shepard, H.M. and Goeddel, D.V. (1981) Nucl. Acids Res. 9, 6153-6166.
54. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
55. Weaver, R. and Weissmann, C. (1979) Nucl. Acids Res. 7, 1175-1193.
56. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980) Cell 20, 555-566.
57. Wilson, A.C., Carlson, S.S. and White, T.J. (1977) Annu. Rev. Biochem. 46, 573-639.
58. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Cell 21, 627-638.
59. Baltimore, D. (1981) Cell 24, 592-594.
60. Egel, R. (1981) Nature 290, 191-192.
61. Dayhoff, M.O. (1978) Atlas of Protein Sequence and Structure, National Biomedical Res. Found., Washington.