

THE PRESENCE OF HOMOLOGOUS REGIONS BETWEEN INTERFERON SEQUENCES

by

JAN TAVERNIER and WALTER FIERIS

Laboratory of Molecular Biology, State University of Ghent,
Ledeganckstraat 35, B-9000 Ghent, Belgium

Dedicated to the 80th birthday of Professor Dr. HEINZ HOLTER

Keywords: Interferons, polypeptide sequences, intervening sequences, 3-dimensional structure

The amino acid sequences derived from cloned interferon genes are compared. Human leukocyte and fibroblast interferons (IFN- α , IFN- β ; both type I interferons) are known to be related (T. TANIGUCHI et al., *Nature* 285, 1980). Comparison of human and murine α - and β -type interferons suggests the presence of two major homology blocks within type I interferon sequences. Also, although immune interferon (IFN- γ ; type II) shows obvious differences in its biological properties and genomic organization, some sequence homology with type I interferons is apparent (W.F. DEGRADO et al., *Nature* 300, 1982; L.B. EPSTEIN, *Nature* 295, 1982). This is especially true for the region 41-104 of the IFN- γ polypeptide which shows homology to both IFN- α and IFN- β . This region is almost completely confined to the third exon of the human IFN- γ gene. When superposed on a tentative 3-dimensional model of type I interferons (M.J.E. STERNBERG & F.E. COHEN, *Int. J. Biol. Macromol.* 4, 1982), a correlation with a domain encompassing two α -helical structures is found. This may have implications for the evolution of the interferon gene family.

1. INTRODUCTION

Interferons can be divided into three distinct classes: α (leukocyte), and β (fibroblast) which are both type I interferons and γ (immune) interferon, a type II interferon. Differences in induction, cellular origin and specificity, as well as the physicochemical and serological properties of these interferons justify this classification.

The application of recombinant DNA-technology has led to the isolation of genes coding for these interferon types from both the human (5,

8, 15, 18, 23, 29) and the murine system (9, 17, 20, 25). In contrast with the unique status of the IFN- β and IFN- γ genes (6, 16, 30, 31), a gene family coding for a whole set of different but related IFN- α polypeptides was found and many of these were shown to be actively transcribed upon induction in humans (1, 14, 22) and in rodents (9, 25). The existence of approximately 20 closely related, but different genes and pseudogenes within the IFN- α multigene family has

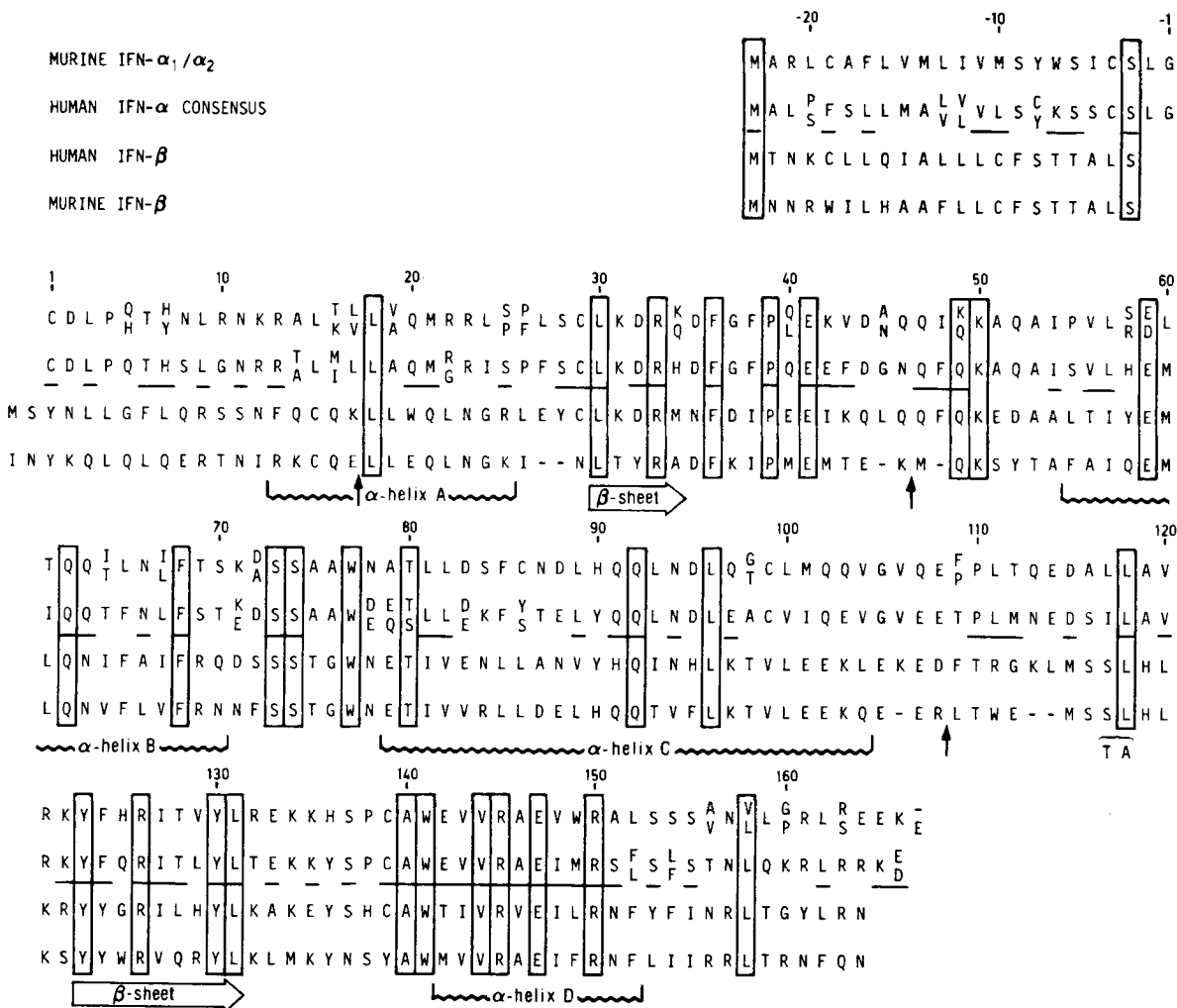


Figure 1. Homology between type I interferons.

Numbering above refers to the human IFN- α polypeptide. Amino acids are given in the one letter notation (3). A human IFN- α consensus sequence (based on ref. 34) is given. Both murine IFN- α_1/α_2 interferons are indicated. The human IFN- α sequence underlining indicates amino acids common to all IFN- α 's. Dashes present in the murine IFN- β polypeptide sequence indicate deletions required for optimal alignment. Only the conserved amino acids present in all interferon sequences are boxed. The location of α -helices and β -sheets as predicted by STERNBERG and COHEN (4) is also shown. Arrows indicate superposition of the INF- γ splice sites, based on the observed homology between type I and type II interferons (see text).

so far been documented in the human genome (34). Although there are some reports on multiple IFN- β genes in humans and mice, these still need further substantiation. Both the human IFN- α multigene family and the human IFN- β gene are located on the short arm of chromosome 9, within the p_{21} - p_{16} region (32). Chromo-

some walking indicated the organization of the IFN- α genes at least partly as tandem arrays (2). The IFN- β gene is not located in the immediate vicinity of the IFN- α gene cluster (19). The mouse IFN- α gene cluster is located on chromosome 4 (21). The human IFN- γ gene has been assigned to the long arm of chromosome 12

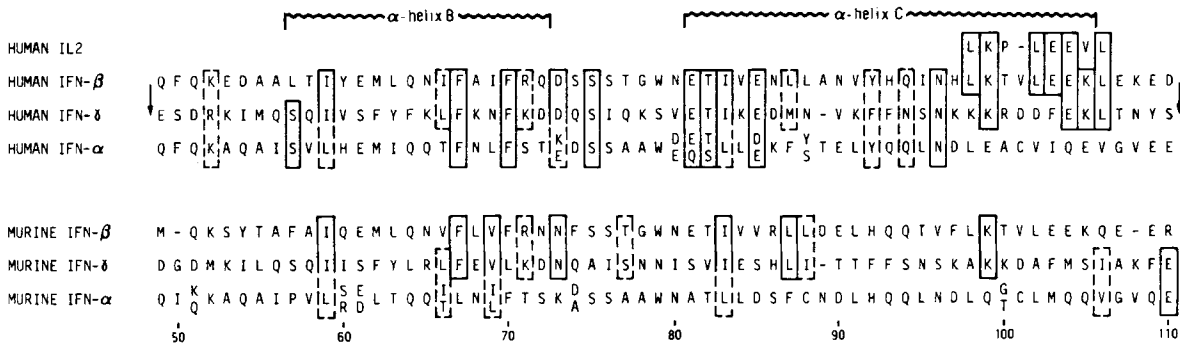


Figure 2. Homology between type I and type II interferons.

Only the region corresponding to the human IFN- γ exon 3 is given. Arrows indicate the splice sites flanking the third exon of the human IFN- γ gene. The homologous regions of the type I interferon genes correspond to residues 47-108 and 49-110 of human IFN- α and IFN- β , respectively. Numbering (underneath) is according to the human IFN- β polypeptide. Full boxes show perfect homology. Dashed boxes indicate conservative changes (see Table I). A dash indicates the absence of an amino acid. The location of the α -helices B and C is also indicated.

(q24.1) (32). All type I interferon genes characterized so far lack intervening sequences (22, 30). The immune interferon gene on the other hand contains three introns in its coding region (16, 31).

2. COMPARISON OF INTERFERON AMINO ACID SEQUENCES

2.1. Comparison of type I interferons

Figure 1 shows a comparison of human and murine type I interferon polypeptide sequences. As a representative amino acid sequence for the multiple human IFN- α polypeptides, an IFN- α consensus sequence, constructed by comparing the two predominant IFN- α subfamilies, is given (34). TANIGUCHI and colleagues (28) indicated the presence of three distinct homology regions by comparing human IFN- α_1 and human IFN- β gene sequences. Within an overall homology of 45% at the nucleotide, and 29% at the amino acid level, two regions stand out in the mature polypeptide; a first region, from position 29-79 (IFN- α numbering) with 41% amino acid homology, and a second from position 116-150 (IFN- α numbering) with 54% amino acid homology. Also, a weak homology was indicated within the signal peptide which obviously is rich in hydrophobic residues. These data are confirmed by extending the comparison of type I interferons to the murine system. As can be

deduced from Figure 1, especially the 118-150 region is strongly conserved: 33% perfect homology between all type I interferon sequences. The conservation of the 30-80 region is somewhat weaker (27%). These two regions might represent two receptor (target) interaction sites (idiotopes) (26).

2.2. Comparison of type I and type II interferon sequences

In contrast with type I interferons, immune interferon is coded by a spliced gene. We, and others (16, 31), showed the presence of three intervening sequences within the coding region of the IFN- γ gene. All three introns are exactly between codons, namely between residues 18/19, 41/42 and 102/103.

Type I and type II interferon polypeptide sequences can be compared when human IFN- β and IFN- γ sequences are aligned as shown in Figure 2. With the introduction of a single gap between IFN- γ residues 81 and 82, a homology of 33% is found between residues 59 and 98 (IFN- γ numbering), or as much as 44% when conservative changes are included (leucine/isoleucine, methionine/leucine, lysine/arginine, phenylalanine/tryptophane, glutamine/asparagine). The homology is clearly restricted to the third exon of the IFN- γ gene (Table I). As can be deduced from Table I, this alignment gives

Table I.
Homology percentages between type I and type II interferons.

Nos. of amino acids	41		61		44 (35)		Total
	Exon 1+2		Exon 3		Exon 4		
human β/γ	(9.8)	12.2%	(22.6)	32.3%	(6.8)	6.8%	(15.1) 19.9%
murine β/γ	(7.3)	9.8%	(11.3)	16.1%	(2.9)	2.9%	(8.0) 11.7%
human α/γ	(4.9)	7.3%	(12.9)	22.6%	(9.1)	11.4%	(10.3) 15.8%
murine α/γ	(2.4)	2.4%	(1.6)	9.7%	(2.9)	2.9%	(2.2) 6.6%

The homology percentages for the amino acid sequences are given for the total polypeptide (right column) and for the regions corresponding to the IFN- γ exons 1+2, exon 3 and exon 4, respectively. The number of amino acids per region is given on top. (The number indicated between parenthesis refers to murine IFN- γ). The homology percentages listed are based on perfect homology positions and on the following conservative changes: leucine/isoleucine, leucine/methionine, isoleucine/valine, arginine/lysine, glutamine/asparagine, glutamic acid/aspartic acid, tryptophane/phenylalanine, and serine/threonine. The percentages of perfect homology are given between parentheses.

analogous results when murine IFN- β and IFN- γ , human IFN- α and IFN- γ , and murine IFN- α and IFN- γ are compared, although the conservation is less pronounced in the murine system.

2.3. Superposition on a type I interferon 3-dimensional model

A secondary and tertiary structure model has been proposed for type I interferons (27). This model shows four α -helices in a right-handed bundle and two short β -sheet structures which are located at one side of the bundle (Figure 3).

The observed type I/type II homology allows us to superimpose the conserved region on this tertiary structure model. Remarkably, the splice sites, flanking the IFN- γ third exon are both located in non-structured regions between the predicted β -sheets and α -helices (Figure 3). The splice sites thus border two α -helices (B and C), one of which (B) was found, in this alignment, to be partly amphiphilic for α , β and γ interferons (4). The helix-B was found to be conserved between type I interferons. With respect to the assumption that the proposed 3-dimensional model for type I interferons also holds for IFN- γ , it should be noted, however, that IFN- γ is a dimer (24, 35).

3. DISCUSSION

The observation that the homology between type I and type II interferons is limited to the

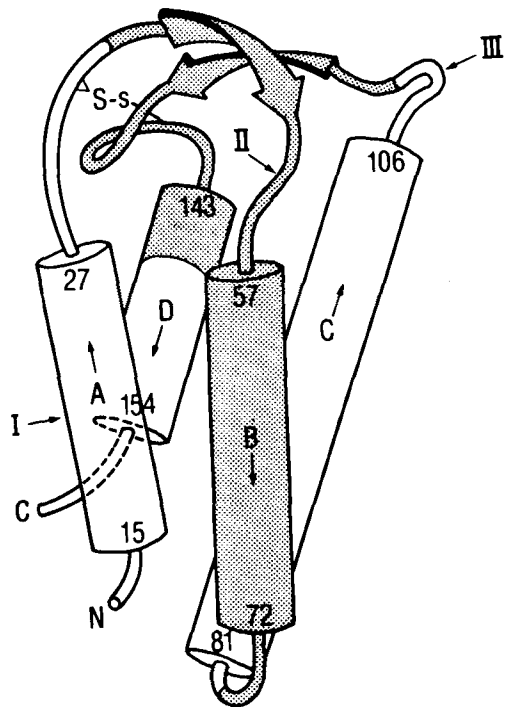


Figure 3. Predicted type I interferon structures (adapted from STERNBERG and COHEN, ref. 27).

Cylinders (A, B, C and D) and arrows represent α -helices and β -sheets, respectively. The numbering is according to the IFN- β polypeptide sequence. The shadowed regions show the homology blocks between type I interferons (Figure 1). Arrows indicate the superimposed IFN- γ splice sites, assuming the 3-D model can also be adapted for this type II interferon (cf. Figure 2).

third exon of the IFN- γ gene is intriguing in view of Gilbert's "domain" theory (12). According to this hypothesis, proteins consist of several functional "domains" separated in higher organisms at the genomic level by intervening sequences. Recombination between introns would then provide a mechanism to exchange "domains" between different proteins ("exon shuffling", ref. 13). This might be reflected at the protein level by the positioning of the splice sites in non-structured regions of the polypeptide, as in fact is illustrated in Figure 3, assuming that the 3-D model proposed for type I interferons can also be adapted for IFN- γ . The third exon of the IFN- γ gene might represent a functional "domain", related to the type I interferons, while the other parts of the IFN- γ sequence might be evolutionarily of different origin. In this perspective, it is interesting to mention also a short homology observed between analogous regions in the human interleukin 2 (which, together with IFN- γ , is the only human lymphokine gene whose structure has been elucidated thus far (7, 11)), and IFN- β genes. Indeed, as seen in Figure 2, the sequence *Leu-Lys-Pro-Leu-Glu-Glu-Val-Leu* (IL2: 83-90) resembles the sequence *Leu-Lys-Thr-Val-Leu-Glu-Glu-Lys-Leu* (IFN- β : 98-106); in this case as well, this homology is located in the third exon of the IL2 gene. The observation that a better homology is obtained between IFN- γ and IFN- β rather than IFN- α might explain some competition for receptor binding observed between the former two polypeptides (33). Whether the proposed alignment between type I and type II interferons is significant remains to be proven. As many of the interferon polypeptides mentioned have been efficiently expressed in prokaryotic and/or eukaryotic systems, large scale preparation has become possible. Hence, crystallization of the proteins may allow deduction of the exact tertiary structures. This will certainly help us to understand the structure-function relationship of these remarkable proteins in molecular detail and this may reveal the significance of conserved domains.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. J. VANDEKERCKHOVE and A. VERSTRAETE for critical

advice. We also thank F. SHAPIRO for editorial, W. DRIJVERS for artistic and L. KIEKENS-DE HENAU for typing assistance.

This research was supported by grants from Biogen, S.A. and from the Geconcerteerde Onderzoeksakties of the Belgian Ministry of Sciences.

REFERENCES

1. ALLEN, G. & K.H. FANTES: A family of structural genes for human lymphoblastoid (leukocyte-type) interferon. *Nature* 287, 408-411 (1980)
2. BRACK, C., S. NAGATA, N. MANTEI & C. WEISSMANN: Molecular analysis of the human interferon- α gene family. *Gene* 15, 379-394 (1981)
3. DAYHOFF, M.O.: Atlas of protein sequence and structure. Natl. Biomed. Res. Found., Washington, D.C. (1978)
4. DEGRADO, W.F., Z.R. WASSERMAN & V. CHOWDRY: Sequence and structural homologies among type I and type II interferons. *Nature* 300, 379-381 (1982)
5. DERYNCK, R., J. CONTENT, E. DECLERQ, G. VOLCKAERT, J. TAVERNIER, R. DEVOS & W. FIERS: Isolation and structure of a human fibroblast interferon gene. *Nature* 285, 542-547 (1980)
6. DEGRAVE, W., R. DERYNCK, J. TAVERNIER, G. HAEGEMAN & W. FIERS: Nucleotide sequence of the chromosomal gene for human fibroblast (β) interferon and of the flanking regions. *Gene* 14, 137-143 (1981)
7. DEGRAVE, W., J. TAVERNIER, F. DUERINCK, G. PLAETINCK, R. DEVOS & W. FIERS: Cloning and structure of the human interleukin-2 chromosomal gene. *The EMBO J.* 2, 2349-2353 (1983)
8. DEVOS, R., H. CHEROUTRE, Y. TAYA, W. DEGRAVE, H. VAN HEUVERS WYN & W. FIERS: Molecular cloning of human immune interferon cDNA and its expression in eukaryotic cells. *Nucl. Acids Res.* 10, 2487-2501 (1982)
9. DYKEMA, R., P. POWWELS, A. DE REUS & H. SCHELLEKENS: Structure and expression in *Escherichia coli* of a cloned rat interferon- α gene. *Nucl. Acids Res.* 12, 1227-1242 (1984)
10. EPSTEIN, L.B.: Interferon-gamma: success, structure and speculation. *Nature* 295, 453-454 (1982)
11. FUJITA, T., C. TAKAOKA, H. MATSUI & T. TANIGUCHI: Structure of the human interleukin-2 gene. *Proc. Natl. Acad. Sci. USA* 80, 7437-7441 (1983)
12. GILBERT, W.: Why genes in pieces? *Nature* 271, 501 (1978)
13. GILBERT, W.: Introns and exons: playground of evolution. In "Eukaryotic Gene Regulation" (Aca-

- demic Press, 1979)
14. GOEDDEL, D.V., D.W. LEUNG, T.J. DULL, M. GROSS, R.M. LAWN, R. MCCANDLESS, P.H. SEEBURG, A. ULLRICH, E. YELVERTON & P.W. GRAY: The structure of eight distinct cloned human leukocyte interferon cDNAs. *Nature* 290, 20-26 (1981)
 15. GOEDDEL, D.V., E. YELVERTON, A. ULLRICH, H.L. HEYNEKER, G. MIOZARRI, W. HOLMES, P.H. SEEBURG, T. DULL, L. MAY, N. STEBBING, R. CREA, S. MAEDA, R. MCCANDLESS, A. SLOMA, J.M. TABOR, M. GROSS, P.C. FAMILLETTI & S. PESTKA: Human leukocyte interferon produced by *E. coli* is biologically active. *Nature* 287, 411-416 (1980)
 16. GRAY, P.W. & D.V. GOEDDEL: Structure of the human immune interferon gene. *Nature* 298, 859-863 (1982)
 17. GRAY, P.W. & D.V. GOEDDEL: Cloning and expression of murine immune interferon cDNA. *Proc. Natl. Acad. Sci. USA* 80, 5842-5846 (1983)
 18. GRAY, P.W., D.W. LEUNG, D. PENNICA, E. YELVERTON, R. NAJARIAN, C.C. SIMONSEN, R. DERYNCK, P.J. SHERWOOD, D.M. WALLACE, S.L. BERGER, A.D. LEVINSON & D.V. GOEDDEL: Expression of human immune interferon cDNA in *E. coli* and monkey cells. *Nature* 295, 503-508 (1982)
 19. GROSS, G., V. MAYR, W. BRUNS, F. GROSVELD, H.H.-M. DAHL & J. COLLINS: The structure of a thirty-six kilobase region of the human chromosome including the fibroblast interferon gene IFN- β . *Nucl. Acids. Res.* 9, 2495-2507 (1981)
 20. HIGASHI, Y., Y. SOKAWA, Y. WATANABE, Y. KAWADE, S. OHNO, C. TAKAOKA & T. TANIGUCHI: Structure and expression of a cloned cDNA for mouse interferon- β . *J. Biol. Chem.* 258, 9522-9529 (1983)
 21. KELLEY, K.A., C.A. KOZAK, F. DANDROY, F. SOR, D. SKUP, J.D. WINDASS, J. DE MAEYER-GUIGNARD, P.M. PITHA & E. DE MAEYER: Mapping of murine interferon- β genes to chromosome 4. *In press* (1984)
 22. NAGATA, S., N. MANTEI & C. WEISSMANN: The structure of one of the eight or more distinct chromosomal genes for human interferon- α . *Nature* 287, 401-408 (1980)
 23. NAGATA, S., H. TAIRA, A. HALL, L. JOHNSRUD, M. STREULI, J. ECSODI, W. BOLL, K. CANTELL & C. WEISSMANN: Synthesis in *E. coli* of a polypeptide with human leukocyte interferon activity. *Nature* 284, 316-320 (1980)
 24. SCAHILL, S.J., R. DEVOS, J. VAN DER HEYDEN & W. FIERS: Expression and characterization of the product of a human immune interferon cDNA gene in Chinese hamster ovary cells. *Proc. Natl. Acad. Sci. USA* 80, 4654-4658 (1983)
 25. SHAW, G.D., W. BOLL, H. TAIRA, N. MANTEI, P. LENGVEL & C. WEISSMANN: Structure and expression of cloned murine IFN- α genes. *Nucl. Acids Res.* 11, 555-573 (1983)
 26. STREULI, M., A. HALL, W. BOLL, W.E. STEWART, S. NAGATA & C. WEISSMANN: Target cell specificity of two species of human interferon- α produced in *E. coli* and of hybrid molecules derived from them. *Proc. Natl. Acad. Sci. USA* 78, 2848-2852 (1981)
 27. STERNBERG, M.J.E. & F.E. COHEN: Interferon: a tertiary structure predicted from amino acid sequences. *Int. J. Biol. Macromol.* 4, 137-144 (1982)
 28. TANIGUCHI, T., N. MANTEI, M. SCHWARZSTEIN, S. NAGATA, M. MURAMATSU & C. WEISSMANN: Human leukocyte and fibroblast interferons are structurally related. *Nature* 285, 547-549 (1980)
 29. TANIGUCHI, T., S. OHNO, Y. FUJII-KURIYAMA & M. MURAMATSU: The nucleotide sequence of human fibroblast interferon cDNA. *Gene* 10, 11-15 (1980)
 30. TAVERNIER, J., R. DERYNCK & W. FIERS: Evidence for a unique fibroblast interferon (IFN- β) chromosomal gene, devoid of intervening sequences. *Nucl. Acids Res.* 9, 461-471 (1981)
 31. TAYA, Y., R. DEVOS, J. TAVERNIER, H. CHEROUTRE, G. ENGLER & W. FIERS: Cloning and structure of the human immune interferon- γ chromosomal gene. *The EMBO J.* 1, 953-958 (1982)
 32. TRENT, J.M., S. OLSON & R.M. LAWN: Chromosomal localization of human leukocyte, fibroblast and immune interferon genes by means of *in situ* hybridization. *Proc. Natl. Acad. Sci. USA* 79, 7809-7813 (1982)
 33. VILCEK, J.: The importance of having gamma. II. Interferon-gamma receptor. I. Gresser, ed., Academic Press, Interferon 4, 129-154 (1982)
 34. WEISSMANN, C., S. NAGATA, W. BOLL, M. FOUNTOULAKIS, A. FUJISAWA, J.-I. FUJISAWA, J. HAYNES, K. HENCO, N. MANTEI, H. RAGG, C. SCHEIN, J. SCHMID, G. SHAW, M. STREULI, H. TAIRA, K. TODOKORO & U. WEIDLE: Structure and expression of human IFN- α genes. *Phil. Trans. R. Soc. London.* B299, 7-28 (1982)
 35. YIP, Y.K., B.S. BARROWCLOUGH, C. URBAN & J. VILCEK: Molecular weight of human interferon is similar to that of other interferons. *Science* 215, 411-413 (1982)