PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

# Purification and Structural Analysis of Interferon [and Discussion]

M. Rubinstein

*Phil. Trans. R. Soc. Lond. B* 1982 **299**, 39-50
doi: 10.1098/rstb.1982.0104

*Phil. Trans. R. Soc. Lond.* B **299**, 39–50 (1982)
*Printed in Great Britain*



# Purification and structural analysis of interferon

By M. Rubinstein
*Department of Virology, Weizmann Institute of Science, P.O. Box 26, Rehovot, Israel*

Studies with crude or partly purified interferon have provided a significant amount of structural information. However, complete biochemical characterization required purification to homogeneity. Earlier work on fractionation has met with many difficulties because interferon was available only in minute quantities. A scale-up of production, adaptation of multi-step purification schemes, use of high-resolution separation techniques and highly sensitive analytical methods have yielded pure interferons and hence many structural data. Specific activities, amino-acid compositions, partial sequences and structural homologies of many interferons were determined. Finally, cloned copy DNA (cDNA) fragments derived from specific interferon mRNA, as well as isolated interferon genes, have been sequenced and the data were used to elucidate complete sequences of many interferons with a high degree of confidence.

## 1. Introduction

Since the discovery of interferon (Isaacs & Lindenmann 1957), much effort has been devoted to its purification as an essential step in its complete chemical and biological characterization. Early work on isolation met with many difficulties and it was soon realized that interferon was a highly potent protein, produced only in minute quantities and obtained in a mixture with many other proteins. Purification to homogeneity of such a protein required many steps of concentration and fractionation, and since losses were bound to occur in each step, a large amount of starting material had to be used. Even if enough crude interferon was available, the resolution obtained by most chromatographic techniques was insufficient. Finally, the amount of pure material that could be obtained was too small for most analytical methods.

Several approaches were used to study the structure of interferon. Even before pure interferon became available, a significant number of structural data were obtained by following the biological activity. In this way molecular masses, the presence or absence of sugar residues and the presence of disulphide bonds were all determined with crude or with partly purified material. Another approach used immunological methods to identify structural relatedness of interferons from different sources. They later served as a basis for the current classification of interferons into three major groups: IFN-α or leucocyte interferon, produced by virus-induced leucocytes; IFN-β or fibroblast interferon, produced by poly(rI)·poly(rC) induced fibroblasts; IFN-γ or immune interferon produced by mitogen-induced leucocytes (Stewart *et al.* 1980*a*). These three groups were found to be antigenically distinct. Still another indirect approach gave by far the largest amount of structural information. This was achieved by use of recombinant DNA techniques. Complete protein sequences were deduced from sequencing of their respective cloned cDNA obtained by reverse transcription of the interferon mRNA. Although DNA sequencing is both faster and more reliable than protein sequencing, direct analysis of the proteins must also be done to confirm unequivocally the indirect data and to find structural features that are determined post-translationally. These include the N-termini of the mature

proteins, other proteolytic cleavages, the location of the disulphide bonds and possible glycosylation sites.

Direct analysis of pure interferon eventually became possible as high-resolution protein separation techniques were developed and used to obtain microgram quantities of pure interferon. Highly sensitive analytical methods were then used for the structural analysis. The procedures developed for the purification of various human interferons and the structural data obtained, either by direct analysis of pure interferon or by indirect methods, are reviewed in this paper.

## 2. Purification of interferon

### (*a*) *Human leucocyte interferon, IFN-α*

Complete purification of many species of IFN-α was obtained by a multi-step procedure involving reverse-phase and normal-phase high-performance liquid chromatography (h.p.l.c.) (Rubinstein *et al.* 1978, 1979, 1981). For the initial studies, leucocytes of normal donors were used as a source of IFN-α. It was later found that leucocytes of chronic myelogenous leukaemia patients produced interferon identical to that obtained from normal donors (Rubinstein *et al.* 1980*b*). Large numbers of leucocytes could be conveniently taken from a single patient by a procedure called leukapheresis. These cells were therefore used for the subsequent isolation of all IFN-α subspecies. H.p.l.c. on a polar bonded silica gave three groups of leucocyte interferons (IFN-α) which were labelled, α, β or γ, according to their order of elution. (Unfortunately, the same notation was assigned later for leucocyte, fibroblast and immune interferon, respectively.) Each group of leucocyte interferons was further resolved by h.p.l.c. on octyl silica. A typical fractionation is shown in figure 1. The isolated species of leucocyte interferon were labelled $\alpha_1$, $\alpha_2$, $\beta_1$, $\beta_2$, $\beta_3$, $\gamma_1$, $\gamma_2$ and $\gamma_3$. Other minor peaks were obtained occasionally. All these species were homogeneous as determined by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-p.a.g.e.) and gave antiviral activity both on human fibroblasts and bovine MDBK cells. In addition, they all inhibited the proliferation of Daudi lymphoblastoid cells (Evinger *et al.* 1980). At least one of the species was tested and found to augment Natural Killer (NK) cell activity, showing for the first time that this property is related to interferon itself and not to other factors that could be present in crude interferon (Ortaldo *et al.* 1980).

The absolute specific activity of pure IFN-α was determined to be $2 \times 10^8$ units $mg^{-1}$.† This value was later confirmed by several independent studies. Claims of a higher specific activity, as much as $2 \times 10^9$ u $mg^{-1}$ were made, but were not supported by a reliable protein determination, which has been quite difficult with the limited amounts of interferon available.

Human leucocyte interferon from the lymphoblastoid Namalwa cells was purified to homogeneity by a multi-step procedure that included antibody affinity chromatography and preparative polyacrylamide gel electrophoresis (Zoon *et al.* 1979). The pure species of interferon had a specific activity of $2.5 \times 10^8$ u $mg^{-1}$. Another multi-step fractionation procedure which included gel filtration, Cu-chelate, blue dextran and antibody chromatography followed by gradient SDS-p.a.g.e. was used to obtain two major and three minor species of IFN-α (Berg & Heron 1980).

Monoclonal antibodies for IFN-α were developed and used for purification (Secher & Burke 1980; Staehelin *et al.* 1981*a*; Montagnier *et al.* 1980). Radiolabelled interferon was shown to

† Based on an interferon standard supplied by The National Institutes of Health, Bethesda, Maryland, U.S.A.


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

bind specifically to the Sepharose-bound monoclonal antibody and a 5000-fold purification in a single step was demonstrated. The purified interferon had a specific activity of $1.8 \times 10^8$ u mg$^{-1}$. Monoclonal antibodies have recently been used on a large scale for purification of bacterially produced IFN-α (Staehelin *et al.* 1981*b*). It appears that this is the simplest and most reliable purification method for α-interferons. In view of the heterogeneity of IFN-α, it is unlikely that

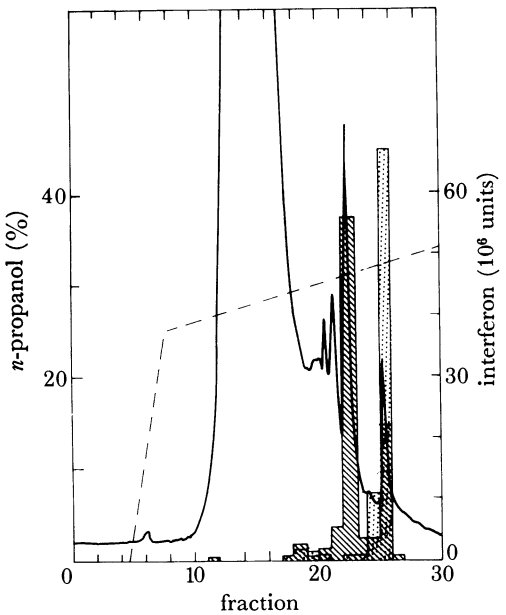


FIGURE 1. Reversed-phase h.p.l.c. of human IFN-α. Partly purified IFN-α was loaded on a Lichrosorb RP-8 column (E. Merck, 4 × 250 mm), equilibrated in a 1 M pyridine – 2 M formic acid buffer and eluted with a gradient of *n*-propanol (broken line) at a flow rate of 0.25 ml min$^{-1}$. Biological activity was measured with bovine MDBK cells (▨) and human FS-11 cells (⊡). Protein (solid line) was monitored by an automated fluorescamine analyser.

a single monoclonal antibody will bind all the subspecies with the same avidity. A battery of monoclonal antibodies will probably be required for the various subspecies. Nevertheless, the structural homology of the various IFN-α subspecies will cause cross-reactivity and will make it very difficult, if not impossible, to produce antibodies that will react specifically with a single subspecies of IFN-α.

### (*b*) *Human fibroblast interferon, IFN*-β

The first successful purification of fibroblast interferon was achieved by a multi-step procedure that included ammonium sulphate precipitation, gel filtration, cation exchange chromatography and preparative polyacrylamide gel electrophoresis (Knight 1976). The specific activity of the pure material was $2 \times 10^8$ u mg$^{-1}$. A significant improvement in the purification procedure was achieved with the use of Sepharose Blue as a specific adsorbant of IFN-β in 1 M NaCl. Pure interferon was obtained by a step gradient of ethylene glycol (Knight & Fahey 1981). The nature of the interaction between the blue dye and IFN-β is not known. It is interesting to note that rare-earth ions such as $La^{3+}$, which stabilized IFN-β (Sedmak & Grossberg 1979), prevented its binding to Sepharose Blue (Friesen *et al.* 1981).

IFN-β was also purified by Con A-Sepharose chromatography (Davey *et al.* 1976), by

Zn-chelate chromatography (Edy *et al.* 1977) and by h.p.l.c. on octyl and diphenyl silica (Friesen *et al.* 1981). Monoclonal antibodies were recently developed for IFN-β (Hochkeppel *et al.* 1981; Nyari *et al.* 1981), and their specificity was demonstrated with the aid of $^{125}$I-labelled IFN-β. As with most other purification systems, monoclonal antibody columns will become the method of choice for IFN-β purification.

### (*c*) *Human immune interferon, IFN-γ*

Purification of IFN-γ has been more difficult than that of IFN-α and β because the yields of crude IFN-γ are lower and because IFN-γ is less stable. Significant purification was achieved by specific adsorbtion to controlled pore glass (Georgiades *et al.* 1980), production of IFN-γ in a serum-free medium, chromatography on Con A-Sepharose and gel filtration (Yip *et al.* 1981). A monoclonal antibody to IFN-γ has recently been observed (H. K. Hochkeppel, unpublished). The use of antibody affinity chromatography for purification will require conditions for elution that will not inactivate IFN-γ or the antibody; in this case, a rather stringent demand. The fact that unlike IFN-α and β the biological activity of IFN-γ was almost completely lost on treatment with SDS prevented the use of preparative SDS-p.a.g.e. for efficient purification of IFN-γ, and even the purest purification did not exhibit bands in SDS-p.a.g.e. that could be unequivocally associated with IFN-γ. On scale-up, it was recently demonstrated that about 10 % of the biological activity of IFN-γ was preserved in SDS. Gel electrophoresis of this residual activity gave two major bands of biological activity corresponding to relative molecular masses of 20000 and 25000. A minor band with an apparent relative molecular mass of 50000 was also observed. Preparative SDS-p.a.g.e. was used to obtain the 20000 component in a pure state (Yip *et al.* 1982), but neither the specific activity nor the amino acid composition has yet been reported.

## 3. The structure of interferon

### (*a*) *Human leucocyte interferon*

Human leucocyte interferon is actually a group of structurally related proteins. Each is made of a single non-glycosylated polypeptide with 165–166 amino acid residues and a calculated relative molecular mass of 19500–20000. The heterogeneity of leucocyte interferon was observed on various chromatographic systems including h.p.l.c. (Rubinstein *et al.* 1979, 1981), and on SDS-p.a.g.e. (Stewart & Desmyter 1975). Glycosylation was considered as the most likely explanation for the heterogeneity of IFN-α. Numerous publications define interferons as glycoproteins, and some investigators even used glycosidases or periodate treatment in an attempt to reduce the heterogeneity (Stewart 1979). It should be noted that IFN-α could never be bound satisfactorily to lectin columns (Davey *et al.* 1976). Indeed, amino-sugar analysis of the isolated subspecies of IFN-α, including one of apparently high molecular mass, ruled out the presence of an *N*-glycosidic linkage in IFN-α (Rubinstein *et al.* 1981). The same result was obtained with a partly resolved mixture of IFN-α's from Namalwa cells (Allen & Fantes 1980).

The finding of a multigene family of IFN-α's (Nagata *et al.* 1980) and the finding of unique peptide maps for isolated species of IFN-α (Rubinstein *et al.* 1981) gave the final explanation for the heterogeneity. The glycosylation sequence (Asn-X-Ser or Asn-X-Thr) was absent from all but one species of IFN-α. The amino acid composition was obtained with less than 1 μg of

pure IFN-α by a highly sensitive amino acid analyser based on fluorescamine (Rubinstein *et al.* 1979). The results were later confirmed by comparison with those predicted from the sequences of the cDNAs (Nagata *et al.* 1980). The high abundance of leucine, isoleucine and phenylalanine was consistent with the rather hydrophobic nature of IFN-α. The various species of IFN-α that were isolated were compared in terms of apparent molecular mass, amino acid composition

TABLE 1. AMINO ACID COMPOSITION OF VARIOUS HUMAN INTERFERONS

| | leucocyte interferon type | | | | | | | | | fibroblast | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| amino acid | $A^{(1)}$ | $\beta_1^{(2)}$ | $B^{(1)}$ | $\beta_3^{(2)}$ | $D^{(1,3)}$ | $\gamma_3^{(2)}$ | $F^{(1)}$ | $Ly^{(4)}$ | $\gamma_2^{(2)}$ | calculated[5] | found[6] |
| Asx | 12 | 13 | 16 | 17 | 17 | 17 | 13 | 15 | 14 | 19 | 15 |
| Thr | 10 | 10 | 6 | 9 | 9 | 9 | 8 | 8 | 10 | 7 | 8 |
| Ser | 14 | 12 | 15 | 13 | 13 | 11 | 13 | 11 | 8 | 9 | 7 |
| Glx | 27 | 24 | 26 | 26 | 24 | 25 | 28 | 27 | 27 | 22 | 19 |
| Pro | 5 | 6 | 6 | 5 | 7 | 5 | 5 | 9 | 5 | 1 | 3 |
| Gly | 5 | 6 | 2 | 3 | 3 | 3 | 5 | 11 | 5 | 6 | 6 |
| Ala | 8 | 9 | 9 | 10 | 8 | 10 | 9 | 11 | 9 | 6 | 10 |
| Cys | 4 | 4 | 4 | 4 | 5 | 3 | 5 | 3 | 3 | 3 | 3 |
| Val | 7 | 7 | 7 | 7 | 7 | 6 | 8 | 8 | 8 | 5 | 7 |
| Met | 5 | 6 | 4 | 5 | 6 | 6 | 5 | 4 | 4 | 4 | 4 |
| Ile | 8 | 8 | 10 | 9 | 7 | 7 | 9 | 7 | 9 | 11 | 8 |
| Leu | 21 | 21 | 22 | 22 | 22 | 23 | 18 | 18 | 22 | 24 | 22 |
| Tyr | 5 | 5 | 5 | 5 | 4 | 4 | 3 | 4 | 5 | 10 | 9 |
| Phe | 10 | 10 | 10 | 9 | 8 | 8 | 11 | 7 | 10 | 9 | 8 |
| His | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 5 | 7 |
| Lys | 10 | 10 | 10 | 8 | 8 | 8 | 11 | 10 | 11 | 11 | 11 |
| Arg | 9 | 9 | 10 | 10 | 12 | 11 | 10 | 10 | 9 | 11 | 10 |
| Trp | 2 | n.d.† | 1 | n.d. | 2 | n.d. | 2 | 0.6 | 0.7 | 3 | 3 |
| $GlucNH_2$ | | — | | — | | — | | | — | | 3 |

*References*: (1) Goeddel *et al.* (1980), (2) Rubinstein *et al.* (1981), (3) Mantei *et al.* (1980), (4) Zoon *et al.* (1979), (5) Taniguchi *et al.* (1980*b*), (6) Friesen *et al.* (1981).

† N.d., not determined.

(table 1), peptide mapping (table 2) and in some cases, sequence analysis. It was found that species $\alpha_1$, $\alpha_2$, $\beta_1$ and $\beta_2$ gave indistinguishable amino acid compositions, molecular masses and peptide maps. It was therefore concluded that these subspecies were closely related or even identical in primary structures. All other species had unique structures (table 2). The identity of $\alpha_1$, $\alpha_2$ and $\beta_1$ was also confirmed by sequence analysis of the tryptic peptides (Levy *et al.* 1981). The properties of the various subspecies are summarized in table 3.

A highly sensitive and sophisticated methodology was developed (Hunkapiller & Hood 1978) for automatic Edman sequencing of picomole quantities of proteins. The first 32 amino acid residues of lymphoblastoid IFN-α were determined by this method (Zoon *et al.* 1980). The first 22 residues of a subspecies of IFN-α were similarly determined (Levy *et al.* 1980). The two sequences were identical and it is noteworthy that both have serine as the N-terminus. In retrospect, this was incorrect as the actual N-terminus was cysteine but it could not be identified because of the disulphide bond. Serine is a typical contaminant in the first sequencing cycle. Additional sequence data were obtained by manual sequencing of interferon tryptic peptides (Allen & Fantes 1980; Levy *et al.* 1981). The data obtained matched well with the sequences predicted from the cDNA structures.

The largest amount of structural information was obtained by recombinant DNA methods.



PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 2. Tryptic peptide analysis of IFN-α's by h.p.l.c.

| peptide[1] | retention time (min)[2] | occurrence in species | | | | | position in species $\alpha_1\beta_1^{(1)}$ |
|---|---|---|---|---|---|---|---|
| | | $\alpha_1\alpha_2\beta_1\beta_2$ | $\beta_3$ | $\gamma_1$ | $\gamma_2$ | $\gamma_3$ | |
| | 22.5 | . | ++ | . | . | . | . |
| | 23.5 | . | ++ | . | . | . | . |
| T1 | 33.5–37.5 | ++ | ++ | ++ | ++ | ++ | 145–149 |
| T2 | 36–39 | ++ | ++ | ++ | ++ | . | 121–125 |
| T3* | 60 | (++) | . | . | . | . | 150–155 |
| | 64–65 | . | + | . | . | + | . |
| T4 | 67–68.5 | ++ | . | + | + | . | 113–120 |
| | 77 | . | + | . | . | . | . |
| T6 | 77–79.5 | + | + | + | + | + | 134–144 |
| | 82 | . | . | . | . | + | . |
| T7 | 83–84.5 | ++ | ++ | ++ | ++ | ++ | 126–131 |
| T8 | 89.5–91.5 | + | + | + | + | . | 32–49 |
| | 92–95 | + | . | + | + | + | . |
| | 93–96 | + | . | + | + | + | . |
| | 98.5 | . | . | . | . | + | . |
| | 100–102 | . | . | + | + | . | . |
| T9 | 104–106.5 | ++ | ++ | ++ | ++ | ++ | 14–22 |
| | 107 | . | . | . | . | ++ | . |
| | 111 | . | + | . | . | . | . |
| T10 | 120–123 | + | . | + | + | . | 24–31 |
| | 165–166 | . | ++ | . | + | . | . |
| T11 | 169–172 | + | . | + | + | . | 71–112 |
| T12 | 176–177 | ++ | . | + | . | + | 50–70 |
| | 179 | . | . | ++ | . | . | . |
| | 184 | . | . | . | . | + | . |

+, Present; ++, present, a major peak; (++), occasionally found but as a major peak; $\alpha_1$, $\alpha_1'$, $\beta_1$ and $\beta_2$ gave essentially the same peptide map.

*References:* (1) Levy *et al.* (1981), (2) Rubinstein *et al.* (1981).

* The C-terminal peptide.

Table 3. Properties of human leucocyte interferons

(Rubinstein *et al.* (1981).)

| species | apparent molecular mass | specific activity on bovine MDBK cells ($10^8$ units $mg^{-1}$) | specific activity on human AG-1732 cells ($10^8$ units $mg^{-1}$) |
|---|---|---|---|
| $\alpha_1$ | 18500 | 2.6 | 2.6 |
| $\alpha_2$ | 18500 | 4 | 3 |
| $\beta_1$ | 18500 | 3.4 | 4.4 |
| $\beta_2$ | 18500 | 4 | 2 |
| $\beta_3$ | 26000 | 4 | 3 |
| $\gamma_1$ | 19500 | 2.6 | 2 |
| $\gamma_2$ | 19500 | 4 | 1.5 |
| $\gamma_3$ | 18500 | 3.5 | 0.15 |

† Determined with lysozyme, soybean trypsin inhibitors and carbonic anhydrase as molecular mass markers.

In fact, the entire sequence of several species of IFN-α was deduced from their respective cDNAs (Mantei *et al.* 1980; Nagata *et al.* 1980; Streuli *et al.* 1980; Goeddel *et al.* 1980). In a very comprehensive study, eight sequences were reported (Goeddel *et al.* 1980). These were labelled A–H (figure 2). Additional sequences were found later and it was estimated that there were up to 20 different genes of IFN-α. It was possible to correlate some of the isolated IFN-α

subspecies with the cDNA structures. Such a correlation could be very easy if both sequences were available. However, in few cases, other parameters such as amino acid compositions and unique biological or physiochemical properties could be used for correlation. Based on sequence homology, recombinant IFN-A was found to be equivalent to subspecies $\alpha_1$, $\alpha_2$, $\beta_1$ and $\beta_2$. Based on a unique apparently high molecular mass, a closely matched amino acid composition (table 1) and peptide mapping, recombinant IFN-B was related to the subspecies $\beta_3$. The high activity on the bovine cells, the rather low activity on human cells, peptide mapping and the amino acid compositions could be used to establish the identity of recombinant IFN-D (or $\alpha 1$ according to Weissmann) with subspecies $\gamma_3$. Subspecies IFN-F may very well be the major component of lymphoblastoid interferon as the amino acid compositions and the N-terminal sequences match very well. All in all, about 80% homology was found among the various subspecies of IFN-$\alpha$, and 99 out of 166 residues were identical in all sequences (excluding the pseudogene E). Highly conserved regions were found in positions 23–36; 62–75; 51–97; 115–131 and 133–153. Within these regions only IFN-D (and IFN-E) showed some variations. IFN-D ($\alpha 1$ according to Weissmann) is unique because it is fully active on the bovine MDBK cell line, but the specific activity is much lower on human cells. Studies of recombinant interferon molecules have shown that this activity is associated with the N-terminal portion of the molecule. It is interesting to note that within the N-terminal portions, IFN-D has unique amino acid residues in areas of a rather high conservation; these include: Asp instead of Asn at position 5, Asp instead of Gly at position 10, Ser instead of Phe at position 27, Met instead of Lys at position 31 and Leu instead of Met at position 60. The changes at positions 5 and 10 involve an increase in the negative charge, while a positive charge is replaced by a neutral group at position 31. Such changes are significant and may affect the affinity for the interferon receptor either directly or indirectly by altering the tertiary structure. It was found that two out of thirteen monoclonal antibodies developed against IFN-$\alpha$ did not bind subspecies $\gamma_3$ at all, while all other subspecies were effectively bound in the solid-phase antibody-binding assay (Staehelin *et al.* 1981*a*). This finding supports the theory of a unique structure for $\gamma_3$ which is equivalent to IFN-D. This subspecies is a major component in crude interferon preparations and since its antiviral activity is rather low, it would not be surprising to find that it fulfils a rather special physiological role.

The positions of the disulphide bonds in IFN-A produced in *Escherichia coli* have been determined (Wetzel 1981). For this purpose, interferon was digested with trypsin and the resulting peptides were resolved on reversed-phase h.p.l.c. before and after reduction with a thiol. Those peaks that changed their retention times were isolated and identified by amino acid analysis. It was found by this procedure that $Cys^{1}$ is connected to $Cys^{98}$ and $Cys^{29}$ to $Cys^{138}$. By analogy, it has been assumed that in species IFN-C to IFN-H, $Cys^{1}$ is connected to $Cys^{99}$, while in IFN-B, $Cys^{1}$ is connected to $Cys^{100}$. An additional free Cys was found in position 86 of IFN-D.

The calculated relative molecular mass of all the species was about 20 000 and the number of amino acid residues was 166 (165 in IFN-A with a deletion in position 44). The uniform length of the various subspecies came as a surprise since significant size heterogeneity was predicted from SDS-p.a.g.e. (Stewart & Desmyter 1975). Both low and high molecular masses were observed. Species with a relative molecular mass lower than 20 000 could be produced either by (specific) processing or by (non-specific) proteolysis. Sequence analysis and peptide mapping of the low molecular mass subspecies $\alpha_1$, $\alpha_2$, $\beta_1$ and $\beta_2$ (all related to the cDNA sequence of IFN-A) have shown that they were shorter by 10 amino acids at the C-termini than the cDNA


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(Levy *et al.* 1981). The suggested C-terminus could not be produced by a common endopeptidase.

While proteolysis or another means of cleavage could account for shorter subspecies, relative molecular masses higher than 20000, as in subspecies $\beta_3$, were more difficult to account for. The unusual behaviour of polyacrylamide gels has been described before with other proteins

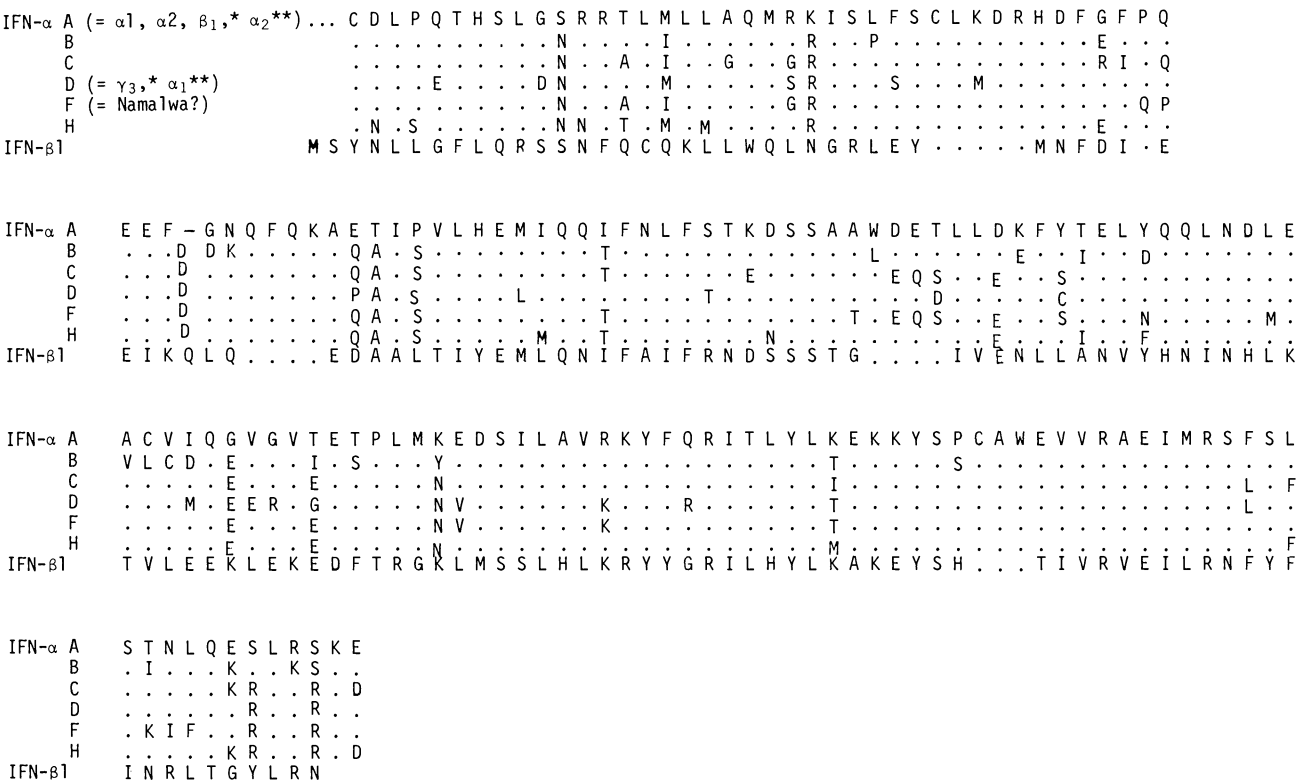


FIGURE 2. Amino acid sequences of various IFNs α and β. A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine. * According to Rubinstein *et al.* (1981); ** according to Mantei *et al.* (1980). IFN-α sequences A–H are from Goeddel *et al.* (1980); IFN-β sequence is from Taniguchi *et al.* (1980*b*).

and might offer the simplest explanation for the 'high' molecular mass species of IFN-α. A study with IFN-D (α1 according to Weissmann) revealed that treatment with SDS and β-mercaptoethanol increased the apparent relative molecular mass from 17500 to 24000 (Stewart *et al.* 1980*b*). However, similar studies with native IFN-α subspecies, including the one equivalent to Weissmann's α1, failed to show any apparent increase of molecular mass upon reduction with β-mercaptoethanol (Rubinstein *et al.* 1981). In addition, subspecies $\beta_3$ (equivalent to recombinant IFN-B) gave a relative molecular mass higher than 20000 regardless of its oxidation state.

### (*b*) *Human fibroblast interferon*

Unlike leucocyte interferon, human fibroblast interferon behaves as a single protein, although additional minor components have been observed occasionally. A 40 kDa band was detected on SDS-p.a.g.e. and was proved to be the dimer of IFN-β, which normally migrates as a


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

20 kDa band. Smaller bands of 17–18 kDa and even 10 kDa were reported as well (Friesen *et al.* 1981). IFN-β is less stable than IFN-α; its biological activity is significantly reduced by oxidation, adsorbance to walls at low protein concentrations or aggregation at high protein concentrations. It has been stabilized by the addition of detergents, carrier proteins such as human serum albumin, antioxidants or rare-earth salts (Sedmak & Grossberg 1979). Amino acid analyses confirmed that IFN-β is a highly hydrophobic protein (Tan *et al.* 1979; Knight *et al.* 1980; Friesen *et al.* 1981). Four lines of evidence indicate that fibroblast interferon is a glycoprotein. It binds to Con A-Sepharose and can be eluted specifically by α-methyl mannoside (Davey *et al.* 1976); pure IFN-β can be stained on polyacrylamide gel by Schiff base (Knight 1976); amino sugar analysis has detected several residues of glucosamine (Tan *et al.* 1979; Friesen *et al.* 1981); finally, a glycosylation signal sequence Asn-Asp-Ser was found in positions 72–74 of mature IFN-β.

The N-terminal sequence of pure IFN-β has been determined by microsequencing methods (Knight *et al.* 1980; Okamura *et al.* 1980; Friesen *et al.* 1981), whereas the complete amino acid sequence has been deduced from the structure of the cDNA (Taniguchi *et al.* 1980*b*; Derynck *et al.* 1980; Houghton *et al.* 1980). All sequence data were identical and have indicated that there is a single, highly conserved gene for IFN-β1 (figure 2); it was compared with that of IFN-α (Taniguchi *et al.* 1980*a*) and 29% homology at the amino acid level was found. Significant homology was observed in positions equivalent to 29–33 and 121–164 of the IFN-α series. These positions are highly conserved within the IFN-α's as well. Two of the three tryptophans and two of the three cysteines of IFN-β were present in the same positions as those of IFN-α. Based on the extent of homology, it was concluded that a common ancestral gene existed some 500–1000 Ma ago, which was the time when the vertebrates first appeared.

IFN-β contains one disulphide bridge and one free cysteine. Based on sequence homology between IFN-α and β, the disulphide bond was predicted to be between residues 31 and 141 (Wetzel 1981). In view of the rather low number of disulphide bonds both in IFN-α and β, the relatively high stability of these interferons is rather surprising. IFN-β, however, is significantly less stable than IFN-α, especially at a neutral pH. Its inactivation is less significant at low pH, suggesting that the inactivation proceeds by oxidation of the –SH group. Aggregation and dimer formation (Knight & Fahey 1981; Friesen *et al.* 1981) represented another facet of the same inactivation process.

### (*c*) *Human immune interferon*

Human immune interferon is different from both IFN-α and β in several ways: it is not stable at pH 2; it induces the antiviral state much more slowly than IFN-α or β; it is probably a more potent inhibitor of cell growth and a more potent activator of Natural Killer cells; unlike IFN-α and β, it does not displace radiolabelled IFN-α from its receptor, and hence, its receptor is different from that of IFN-α and β (Branca & Baglioni 1981). Finally, IFN-γ is synergistic with IFN-α or β in eliciting the antiviral state. All these observations indicate that IFN-γ may have a unique structure with little or no homology with other interferons. A previous nomenclature defining IFN-α and IFN-β as type I and IFN-γ as type II seems to be more sensible than the current nomenclature.

Partially purified IFN-γ exhibits a relative molecular mass of about 50000 on gel filtration columns. A 25000 component was found in Con A induced IFN-γ and was reported to be antigenically similar to IFN-β (Delay *et al.* 1981). IFN-γ did bind to Con A-Sepharose and was


PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

specifically released by α-methyl-*O*-mannoside; hence it is a glycoprotein. It was found to be stable to β-mercaptoethanol at room temperature and to contain free sulphydryl group(s). Isoelectric focusing of PHA-induced IFN-γ has shown that it is a basic protein with a major component of pI 8.6–8.7 and minor components within the pI range 7.5–10 (Yip *et al.* 1981). IFN-γ prepared by induction with staphylococcal enterotoxin A reveals three ranges of pI 4.0–4.3, 4.9–5.3 and 8.4–9.0 (Georgiades *et al.* 1980). Antibodies prepared against this preparation neutralize IFN-γ's prepared with either Con A or PHA (Langford *et al.* 1981). It was therefore suggested that the various inducers of IFN-γ evoke the production of the same or closely related species of IFN-γ.

Molecular cloning of IFN-γ mRNA and expression of the biologically active protein have been reported recently (Gray *et al.* 1982). Only one species of IFN-γ was found; it was 146 amino acid long and behaved similarly to the native IFN-γ. The complete protein sequence was elucidated and it revealed areas rich in basic amino acids such as Lys-Lys-Lys-Arg and Lys-Arg-Lys-Arg, in agreement with the high pI previously reported (Yip *et al.* 1981). There was no sequence homology with type I interferons (IFN-α and β).

## References


Allen, G. & Fantes, K. H. 1980 *Nature, Lond.* **287**, 408–411.

Berg, K. & Heron, I. 1980 *J. gen. Virol.* **50**, 441–446.

Branca, A. & Baglioni, C. 1981 *Nature, Lond.* **294**, 768–770.

Davey, M. W., Sulkowski, E. & Carter, W. A. 1976 *Biochemistry, Wash.* **15**, 704–713.

Delay, M., Billiau, A. & DeSomer, P. 1981 Presented at 5th International Congress on Virology, Strasbourg, France.

Derynck, R., Content, J., DeClercq, E., Volckaer, G., Tavernier, J., Devos, R. & Fiers, W. 1980 *Nature, Lond.* **285**, 542–547.

Edy, V. G., Billiau, A. & DeSomer, P. 1977 *J. biol. Chem.* **252**, 5934–5935.

Evinger, M., Rubinstein, M. & Pestka, S. 1980 In *Interferon: properties and clinical uses* (ed. A. Khan, N. O. Hill & G. L. Dorn), pp. 249–263. Dallas: Leland Fikes Foundation Press.

Friesen, H. J., Stein, S., Evinger, M., Familletti, F. C., Moschera, J., Meienhoffer, S., Shively, J. & Pestka, S. 1981 *Arch. Biochem. Biophys.* **206**, 432–450.

Georgiades, J. A., Langford, M. P., Goldstein, L. D., Blalock, J. E. & Johnson, H. M. 1980 In *Interferon: properties and clinical uses* (ed. A. Khan, N. O. Hill & G. L. Dorn), pp. 97–108. Dallas: Leland Fikes Foundation Press.

Goeddel, D. V., Yelverton, E., Ullrich, A., Heyneker, H. L., Miozzari, G., Holmes, W., Seeburg, P. H., Dull, T., May, L., Stebbing, N., Crea, R., Maeda, S., McCandliss, R., Sloma, A., Tabor, J. M., Gross, M., Familletti, P. C. & Pestka, S. 1980 *Nature, Lond.* **287**, 411–416.

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. & Goeddel, D. V. 1982 *Nature, Lond.* **295**, 503–508.

Hochkeppel, H. K., Menge, U. & Collins, J. 1981 *Nature, Lond.* **291**, 500–501.

Houghton, M., Eaton, M. A. W., Stewart, A. G., Smith, J. C., Doel, S. M., Catlin, G. H., Lewis, H. M., Patel, T. P., Emtage, J. S. & Carey, N. H. 1980 *Nucl. Acid Res.* **8**, 2885–2894.

Hunkapiller, M. W., & Hood, L. E. 1978 *Biochemistry, Wash.* **17**, 2124–2133.

Isaacs, A. & Lindenmann, J. 1957 *Proc. R. Soc. Lond.* B **147**, 258–267.

Knight, E., Jr. 1976 *Proc. natn. Acad. Sci. U.S.A.* **73**, 520–523.

Knight, E., Jr & Fahey, D. 1981 *J. biol. Chem.* **256**, 3609–3611.

Knight, E., Jr, Hunkapiller, M. W., Korant, B. D., Hardy, R. W. F. & Hood, L. E. 1980 *Science, Wash.* **207**, 525–526.

Langford, M. P., Weigent, D. A., Georgiades, J. A., Johnson, H. M. & Stanton, J. 1981 *J. Immunol.* **126**, 1620–1623.

Levy, W. P., Rubinstein, M., Shively, J., Del Valle, U., Lay, C. Y., Moschera, J., Brink, L., Gerber, L., Stein, S. & Petska, S. 1981 *Proc. natn. Acad. Sci. U.S.A.* **78**, 6186–6190.

Levy, W. P., Shively, J., Rubinstein, M., Del Valle, U. & Pestka, S. 1980 *Proc. natn. Acad. Sci. U.S.A.* **77**, 5102–5104.

Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. & Weissmann, C. 1980 *Gene* **10**, 1–10.




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Montagnier, L., Laurent, A. G. & Greust, J. 1980 *C.r. hebd. Seanc. Acad. Sci., Paris* D **291**, 893–896.
Nagata, S., Mantei, N. & Weissmann, C. 1980 *Nature, Lond.* **287**, 401–408.
Nyari, L. J., Tan, Y. H. & Erlich, H. A. 1981 In *The biology of the interferon system* (ed. E. DeMaeyer, G. Galasso & H. Schellekens), pp. 67–71. Amsterdam: Elsevier/North-Holland Biomedical Press.
Okamura, H., Berthold, W., Hood, L., Hunkapiller, M., Inoue, M., Smithson, H. & Tan, Y. H. 1980 *Biochem. biophys. Res. Commun.* **19**, 3831–3835.
Ortaldo, J. R., Pestka, S., Slease, R. B., Rubinstein, M. & Herberman, R. B. 1980 *Scand. J. Immun.* **12**, 365–369.
Rubinstein, M., Rubinstein, S., Familletti, P. C., Gross, M. S., Miller, R. S., Waldman, A. A. & Pestka, S. 1978 *Science, Wash.* **202**, 1289–1290.
Rubinstein, M., Rubinstein, S., Familletti, P. C., Gutterman, J., Hester, J. & Pestka, S. 1980*a* In *Interferon: properties and clinical uses* (ed. A. Khan, N. O. Hill & G. L. Dorn), pp. 45–55. Dallas: Leland Fikes Foundation Press.
Rubinstein, M., Rubinstein, S., Familletti, P. C., Hershberg, R. D., Brink, L. D., Gutterman, J., Hester, J. & Pestka, S. 1980*b* In *Proteins: structure and biological functions* (ed. E. Gross & J. Meienhofer), pp. 99–103. Rockford, Ill.: Pierce Chemical Company.
Rubinstein, M., Rubinstein, S., Familletti, P. C., Miller, R. S., Waldman, A. A. & Pestka, S. 1979 *Proc. natn. Acad. Sci. U.S.A.* **76**, 640–644.
Rubinstein, M., Levy, W. P., Moschera, J. A., Lai, C.-Y., Hershberg, R. D., Bartlett, R. T. & Pestka, S. 1981 *Arch. Biochem. Biophys.* **210**, 307–318.
Secher, D. S. & Burke, D. C. 1980 *Nature, Lond.* **285**, 446–450.
Sedmak, J. J. & Grossberg, S. E. 1979 *Ann. N.Y. Acad. Sci.* **350**, 593.
Staehelin, T., Durrer, B., Schmidt, J., Takacs, B., Stocker, J., Miggiano, V., Stahli, C., Rubinstein, M., Levy, W. P., Hershberg, R. D. & Pestka, S. 1981*a* *Proc. natn. Acad. Sci. U.S.A.* **78**, 1848–1852.
Staehelin, T., Hobbs, D. S., Kung, H. F., Lavy, C. Y. & Pestka, S. 1981*b* *J. biol. Chem.* **256**, 9750–9754.
Stewart, W. E. II 1979 *The interferon system*, pp. 166–167. New York: Springer-Verlag.
Stewart, W. E. II, Blalock, J. E., Burke, D. C., Chany, C., Dunnick, J. K., Falcoff, E., Friedman, R. M., Galasso, G. J., Joklik, W. K. & Vilcek, J. T. 1980*a* *J. Immunol.* **125**, 2353–2358.
Stewart, W. E. II & Desmyter, J. 1975 *Virology* **67**, 68–78.
Stewart, W. E. II, Sarkar, F. H., Taira, H., Hall, A., Nagata, S. & Weissmann, C. 1980*b* *Gene* **11**, 181–186.
Streuli, M., Nagata, S. & Weissmann, C. 1980 *Science, Wash.* **209**, 1343–1347.
Tan, Y. H., Barakat, F., Berthold, W., Smith Johannessen, H. & Tan, C. 1979 *J. biol. Chem.* **254**, 8067–8073.
Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Muramatsu, M. & Weissmann, C. 1980*a* *Nature, Wash.* **285**, 547–549.
Taniguchi, T., Ohno, S., Fujiikur, Y. & Muramatsu, M. 1980*b* *Gene* **10**, 11–15.
Weissmann, C., Nagata, S., Hall, A., Taira, H., Mantei, N., Streuli, M., Mishina, M., Ecsodi, J. & Boll, W. 1980 *Eur. J. Cell* **22**, 17.
Wetzel, R. 1981 *Nature, Lond.* **289**, 606–607.
Yip, Y. K., Barrowclough, B. S., Urban, C. & Vilček, J. 1982 *Proc. natn. Acad. Sci. U.S.A.* **79**, 1820–1824.
Yip, Y. K., Pang, R. H. L., Urban, C. & Vilcek, J. 1981 *Proc. natn. Acad. Sci. U.S.A.* **78**, 1601–1605.
Zoon, K. C., Smith, M. E., Bridgen, P. J., Anfinsen, C. B., Hunkapiller, M. W. & Hood, L. E. 1980 *Science, Wash.* **207**, 527–528.
Zoon, K. C., Smith, M. E., Bridgen, P. J., Zur Nedden, D. & Anfinsen, C. B. 1979 *Proc. natn. Acad. Sci. U.S.A.* **76**, 5601–5605.

## *Discussion*

The nomenclature problem was discussed and Professor Weissmann said that at the Squaw Valley meeting he was going to propose that we had separate systems of nomenclature for (*a*) the genes, (*b*) the peptides, the structure of which would depend on processing and glycosylation as well as DNA sequences, and (*c*) biological functioning, for which the old system of type I and II still had advantages.

The origin and structure of the 26 kDa interferon is still obscure. In reply to questions about the truncated IFN-α which had been isolated by Pestka's groups, and which lacked ten amino acids at the C-terminal end, Dr Rubinstein said that he thought that this might be an artefact. In response to a question as to whether one or multiple species of IFN-α were produced by specialized cells in response to a normal virus infection, it was pointed out that multiple species




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

of IFN-α were produced and isolated from cultures of cells such as the lymphoblastoid Namalwa and the myeloblastoid KG-1 cell lines.

In a further discussion on interferon-like peptides in plants it was suggested that these might arise by 'horizontal' transmission of a gene if, as for plant haemoglobin, it emerged that the structure was too close to that of interferon of vertebrates to fit with the hypothesis of its representing a separate line of evolution from a primordial ancestor.



PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES