Chemical synthesis of a human interferon- α_2 gene and its expression in *Escherichia coli*

M.D.Edge, A.R.Greene, G.R.Heathcliffe, V.E.Moore, N.J.Faulkner, R.Camble, N.N.Petter, P.Trueaman, W.Schuch*, J.Hennam*, T.C.Atkinson, C.R.Newton and A.F.Markham

---

ICI Pharmaceuticals Division, Alderley Park, Macclesfield, Cheshire, SK10 4TG and *ICI Corporate Biosciences Group, The Heath, Runcorn, Cheshire, WA7 4QE, UK

Received 7 June 1983; Revised and Accepted 16 August 1983

ABSTRACT

A 511-base pair DNA fragment encoding human interferon- α_2 has been chemically synthesised and expressed from a lac UV5 and a synthetic trp promoter in *Escherichia coli*. The synthesis involved preparation of 68 oligodeoxyribonucleotides and their enzymic ligation. The product expressed from the trp promoter system had high antiviral activity and displayed biological effects similar to those of Namalwa interferon on natural killer cell activity and in a Daudi cell growth inhibition assay. E.coli minicells containing plasmid DNA with the synthetic IFN- α_2 gene under trp promoter control produce a protein with the same electrophoretic mobility as a sample of authentic IFN- α_2 . The protein from E.coli cross-reacts with the monoclonal antibody NK-2 and was readily purified, close to homogeneity, by immunoabsorption chromatography on NK-2 sepharose.

INTRODUCTION

Human leukocyte interferon- α_2 (IFN- α_2 or LeIF-A) is an antiviral protein with a range of other biological activities (1-4). It is one of the most widely studied of the various leukocyte (IFN- α) interferon species in the laboratory and the clinic. The molecule has features in common with other leukocyte interferons, both in its amino acid and gene sequences (1,5,6). However, the various human IFN- α subtypes have distinct antiviral properties in cell cultures (7) and differ in their effects on cell growth and natural killer cell activity (2). The specific amino acids or domains of the proteins responsible for variations in biological activity are not known.

A number of novel interferon molecules which themselves have distinct biological activities (8,9) has been generated, by exploitation of common restriction sites within coding sequences. However, the range of molecules that can be produced directly by this procedure, for structure/activity studies, is limited.

Rapid chemical synthesis of genes and their expression in bacteria can now be accomplished. A general strategy for gene assembly involves synthesis of short overlapping complementary oligonucleotides (10) (10-20

nucleotide units) and their subsequent enzymic ligation to yield double-stranded fragments for cloning. Re-assembly of such genes using oligonucleotide fragments with deliberate sequence alterations should lead to the expression of novel proteins for biological evaluation.

Here we report the synthesis of a gene encoding human IFN- α_2 and construction of plasmids which direct the synthesis in E.coli of a polypeptide with potent biological activities and the same electrophoretic mobility as a sample of authentic IFN- α_2 .

MATERIALS AND METHODS

Enzymes were purchased from Bethesda Research laboratories or Boehringer and used as directed by the suppliers. Radiochemicals were from Amersham International. Starting materials for oligonucleotide synthesis and hplc materials, unless otherwise mentioned, are as previously described (11,12,13). Immunoradiometric assays (IRMA), based on the NK-2 monoclonal antibody, were purchased from Celltech Ltd., Slough and performed according to manufacturers instructions. E.coli strain JA221 was received from J. Carbon (14), strain C600 from N E Murray (15), strain DS410 from P A Meacock (16) and strain PA340 from the E.coli Genetic Stock Centre (17). Plasmid pAT153 was provided by D J Sherratt (18) and plasmid pPM60 was a gift from P A Meacock. Human Namalwa lymphoblastoid IFN of specific activity 4×10^7 U/mg was purchased from the Israel Institute for Biological Research. Authentic IFN- α_2 was a gift from Professor C Weissmann.

GENERAL METHODS

Small scale plasmid DNA preparations for rapid analysis were made by the methods of Birnboim and Doly (19) or Holmes and Quigley (20). DNA restriction fragments were isolated from polyacrylamide gels by electroelution (21). DNA sequencing was performed by the methods of Maxam and Gilbert (22) or Wallace et al (23). In situ colony hybridisations were performed by the Grunstein-Hogness procedure (24). Transformations of JA221, PA340 and C600 were performed as described previously (25) using L-agar plates (26) containing 33 μ g/ml ampicillin (Penbritin, Beecham) for selection.

Chemical Synthesis of Oligonucleotides

Oligonucleotides were synthesised by solid-phase phosphotriester methodology and purified as described previously (11) except that in Partisil 10-SAX ion-exchange hplc 30% or 60% formamide was substituted for aqueous ethanol in elution buffers (27). Oligonucleotide size and purity were checked by electrophoresis on 20% polyacrylamide/7M urea gels after 5'-labelling with

[γ - ^{32}P] ATP and T4 polynucleotide kinase.

Ligation of oligonucleotides to construct the synthetic IFN- α 2 gene

The oligonucleotides shown in Fig. 1 (0.4nmol), with the exceptions of 1,10,11,20,21,30,31,40,41,50,51,60,61 and 68 were phosphorylated to a specific activity of 8.4 Ci/mmol with [γ - ^{32}P] ATP (11). Reaction mixtures were either combined to give the appropriate small sub-fragments and ethanol precipitated or ethanol precipitated then fractionated on 20% polyacrylamide/7M urea gels. Electroeluates were concentrated by butanol extraction and ethanol precipitated. Yields were monitored by Cerenkov counting.

The small sub-fragments, A to M, were constructed by ligating the appropriate mixture of labelled and unlabelled oligonucleotides (Fig.2). Aliquots of the requisite oligonucleotides were combined and lyophilised. The mixtures were suspended in water (50 μ l), heated to 100° for 5 minutes then cooled slowly to 22°C and lyophilised again. They were resuspended in 25 μ l of 74mM Tris-HCl (pH7.6), 6.6mM MgCl₂, 10mM DTT, 2.9mM ATP, 1.6mM β -mercaptoethanol and 1.6 μ g bovine serum albumin with 2 units of T4 DNA ligase. The reaction mixtures were incubated at 30°C for 10 min. then at 16°C for 24h. Aliquots (1%) of the reaction mixtures were examined on a 15% denaturing polyacrylamide gel. The DNA from the remaining mixtures was ethanol precipitated and fractionated on a 15% non-denaturing polyacrylamide gel. The appropriate DNA fragments were electroeluted and ethanol precipitated after concentration by butanol extraction.

Fragments I,II and III were constructed from the small sub-fragments A to L. B to L were individually 5'-phosphorylated and ethanol precipitated. The appropriate sub-fragments were annealed at 45°C and ligated as described previously. The ligation mixtures were extracted with phenol/chloroform then twice with ether and the DNA ethanol precipitated. The large fragments I,II and III were then isolated by electrophoresis on a 10% non-denaturing polyacrylamide gel, electroelution and ethanol precipitation.

The IFN- α 2 gene was constructed from fragments I,II and III (0.30, 0.20 and 0.07 pmol respectively) and the sub-fragment M(0.20pmol). The components were annealed, ligated and the DNA precipitated as described for the large fragments. The largest product at ca 500-base pairs (bp) was isolated from a 6% non-denaturing polyacrylamide gel to give 4.1ng (12fmol) of the synthetic gene. As a final step before ligation to the vector the Bam HI and Sal I ends were 5'-phosphorylated with [γ - ^{32}P] ATP and T4 polynucleotide kinase.

Construction of recombinant plasmids and cloning procedures

Plasmid pPM60 was digested with Bam HI and Sal I and the 3.2Kb fragment isolated. This purified large fragment (1 μ g) plus 4.1ng (12 fmol) of the synthetic IFN- α_2 gene were treated with DNA ligase (1 unit) for 16hours at 14°C in a total reaction volume of 20 μ l. After transformation of E.coli strain PA340 with this mixture, ampicillin resistant colonies were screened by Grunstein-Hogness colony hybridisation (24) using as probe a ³²P-labelled oligonucleotide from within the synthetic sequence (small sub-fragment H labelled by nick translation). Plasmid DNA was isolated from positive colonies and subjected to restriction analysis with Eco RI, Bam HI, Sal I, Hae III and Hpa II.

Growth of bacterial strains and preparation of extracts for IFN assays.

E.coli strains PA340 and C600 were grown at 37°C to a stationary phase in L broth (28). JA221 cultures were grown at 37°C in Spizizen salts (29) as previously described (30). Lysates were purified by precipitation with trichloroacetic acid and size exclusion chromatography on Ultrogel Aca 54 (31) to a final specific activity of 5.4 x 10⁵ U of interferon/mg of protein before determination of growth inhibition and natural killer cell activity. Samples of gel-filtered extracts were further purified by affinity chromatography on a column of NK-2 Sepharose monoclonal antibody (32) (Celltech Ltd, Slough) and characterised by PAGE-SDS (33) with detection by Coomassie blue or supersensitive silver staining (34).

Preparation of E.coli minicells and ³⁵S-labelled proteins

Plasmid DNA was introduced into the minicell-forming E.coli strain DS410. Minicells were isolated and labelled (100 μ l, A⁶⁰⁰2.0) with ³⁵S-methionine essentially as described (8). Aliquots (20 μ l) of labelled minicell preparations were lysed with an equal volume of 125 mM Tris-HCl (pH6.8), 20% glycerol, 2% SDS and 2% β -mercaptoethanol at 100°C for 3 minutes then loaded onto a 15% PAGE-SDS gel. Visualisation was by fluorography (35).

Interferon bioassays

Antiviral (30), growth inhibition (36) and natural killer cell activity (2,36) assays were performed essentially as described elsewhere.

RESULTS

Chemical synthesis of oligonucleotides

The component oligonucleotides for the IFN- α_2 gene were designed on the basis of the derived amino acid sequence of this protein (1). Codons normally found in highly expressed E.coli proteins were used throughout (37)

except where computer analysis revealed that the resultant oligonucleotide sequences might not be amenable to enzymic ligation. Changes with respect to the native sequence (103 in total) are indicated by dots above the coding strand in Fig.1. The gene was divided into the 68 oligonucleotide fragments shown in Fig. 1. These were chemically synthesised, using phosphotriester solid phase methodology, as previously described(11,12,13), deprotected by standard procedures and purified by two stage high-performance liquid chromatography (hplc) (27). Purified oligonucleotides were labelled using [γ - ^{32} P] ATP and polynucleotide kinase. Their size and purity were confirmed by electrophoresis on 20% polyacrylamide/7M urea gels. Oligomers showing unusual hplc profiles or unusual mobility on polyacrylamide gel electrophoresis were sequenced by the Maxam-Gilbert method (22). Several oligonucleotides do migrate faster than expected on gel electrophoresis, but this can usually be accounted for by their base compositions (38). The mobilities of oligonucleotides 13,39 and 40, however, were so exceptional that sequence analysis was carried out. Sequences 39 and 40 were found to be correct. The 3' base (C) of sequence 13 could not be distinguished by the Maxam-Gilbert procedure.[It appears that secondary structure within the oligonucleotide may alter the electrophoretic mobility resulting in lack of resolution from its n-1 (at the 3' end) degradation product,despite the presence of 7M urea]. Oligonucleotide 13 does ligate to 15 in the formation of group C (Fig.2) and sequence determination of the long strand in group C (11+13+15) clearly showed that the C residue derived from the 3' end of 13 was present (data not shown).

Ligation of oligonucleotides to construct the synthetic IFN- α_2 gene.

The strategy for enzymic ligation of the 68 oligonucleotides to generate the IFN- α_2 gene is shown in Fig.2. This involves the synthesis of 13 double-stranded small sub-fragments (A to M). Certain oligonucleotides were kinased using ATP in the presence of a small amount of [γ - ^{32} P] ATP. Since the A to M groups might be useful as intermediates in subsequent synthesis of selectively modified genes their 5'-protruding ends were left unphosphorylated. These intermediates could be quantified using the low specific activity internal labels then aliquots used for further gene synthesis at any later date following labelling to higher specific activity at the 5'-positions. Fragments A and M were not 5'-phosphorylated until gene assembly was complete because of their self-complementary restriction site sticky ends. Fig. 3a shows an autoradiograph of a 15% polyacrylamide/7M urea gel after electrophoresis of 1% aliquots of reaction mixtures generating the

Fig.1. Nucleotide and amino acid sequence of the synthetic human IFN- α_2 gene. Individual oligonucleotides (1-68) are arrowed. Dots above bases in the coding strand indicate changes with respect to the native sequence. Numbers above each line refer to amino acid position.

Fig.2. Ligation of chemically synthesised oligonucleotides to construct the synthetic gene. Dots at the ends of oligonucleotides indicate the positions of 5'-terminal phosphate groups.

small sub-fragments C,A,E,I,K and L (see Materials and Methods). The bands at 30bp and 45bp are clearly demonstrated. The 45bp strand for group I in this particular preparation was produced in low yield. This is also apparent when the remaining reaction mixtures were electrophoresed in a 15% non-denaturing gel (Fig.3b). The bands corresponding to the double-stranded fragments (45+30bp) are generally the slowest migrating products. The faster migrating major product present in the reaction mixture for sub-fragment I presumably has one or more terminal oligonucleotide(s) missing. The appropriate fragments were eluted from the gel and ethanol precipitated. The large fragments I,II and III were prepared in a similar manner from their constituent small sub-fragments and isolated from a 10% polyacrylamide gel.

The complete gene was assembled from the large fragments I, II and III plus the small sub-fragment M. In this ligation reaction, the ratio of fragments (I:II:III:M) was 4.3:2.9:1:2.9 respectively and because they were prepared at different times the specific activities of fragments I and II were about 10 fold greater than those of III and M. The total reaction mixture was separated on a 6% polyacrylamide gel (Fig.3c) with a marker of an IFN- α_1 gene (514bp Bam HI-Sal I fragment from pIFS1 (11) labelled by nick translation). The slowest migrating product corresponded in size to the IFN- α_1 gene marker and was presumed to be the required IFN- α_2 gene. In view of the excesses of fragments I, II and M over III, together with the higher specific activity of I and II, the relative intensities of the bands (Fig.3c) do not reflect the yield of the product gene. In fact, the IFN- α_2 gene was isolated in 17% overall yield (based on the amount of fragment III used) after electroelution and ethanol precipitation. Before ligation to the

Fig.3. Fractionation of the ligation products of oligonucleotides comprising groups C,A,E,I,K and L (lanes 1 to 6) (a) on a 15% polyacrylamide/7M urea gel (b) on a 15% non-denaturing polyacrylamide gel (c) Isolation of the synthetic IFN- α_2 gene on a 6% polyacrylamide gel.

vector DNA for cloning, the gene was 5'-phosphorylated and again ethanol precipitated.

Molecular cloning of the synthetic IFN- α_2 gene

Plasmid pPM60 (39) was used for initial cloning of the synthetic IFN- α_2 gene. The lac UV5 promoter is situated between Eco RI and Bam HI sites in an orientation such that introduction of the synthetic gene between the Bam HI and Sal I sites should place the interferon coding sequence under its transcriptional control (Fig.4).

Two colonies were isolated from half of the E.coli strain PA340 transformation mixture. Analysis of the plasmid DNAs, designated pIFS1101 and pIFS1102 from each of these colonies showed that both produced the expected 511-bp and 610-bp fragments on digestion with Bam HI/Sal I or Eco RI/Sal I respectively. The expected digestion patterns were also obtained with the restriction endonucleases Hae III and Hpa II (data not shown). To confirm the IFN- α_2 gene structure, pIFS1101 was sequenced by the method of Wallace et al. (23). The plasmid was linearised with Eco RI. Oligonucleotides 9,21,24,37,48,51 (Fig.1) together with primers (Sal 1, ref.23 and a 21 mer designed from the sequence of the lac UV5 promoter)

Fig.4. Construction of recombinant plasmids containing the synthetic IFN- α_2 gene under *lac* UV5 and *trp* promoter control.

reading into the gene from either end were used in dideoxy sequencing reactions. The nucleotide sequence so determined corresponded exactly to that shown in Fig. 1. Given the availability of all the oligonucleotides comprising the gene, this sequencing approach is obviously the most attractive method for this case.

Construction of plasmids carrying a synthetic *trp* promoter/IFN- α_2 gene.

The synthetic IFN- α_2 gene was designed so that it could easily be positioned under the transcriptional control of a synthetic *trp* promoter (30) after cutting the gene at the single *Sau* 3A site (GATC) within the codons for Asp 2/Leu 3 (see sequence in italics, Fig. 1). Fig. 4 outlines the

construction of plasmids in which IFN- α_2 is synthesised from the synthetic trp promoter. A 500bp IFN- α_2 fragment (Sau 3A-Sal I) encoding Asp 2 to the termination codon was isolated from the above plasmid pIFS1101. A 97bp Eco RI-Sau 3A promoter fragment, comprising the synthetic trp promoter, the trp LD102 deletion mutant ribosome binding site and codons for methionine and cysteine 1 was isolated from plasmid pIFS201 (30) which contains the trp promoter/IFN- α_1 construction.

The IFN- α_2 (Sau 3A - Sal I) fragment together with the trp promoter (Eco RI-Sau 3A) fragment were ligated to the large fragment (Eco RI-Sal I) from pAT153 to regenerate the IFN- α_2 gene under control of the trp promoter. This approach is possible since the sequences of the first four amino acids of IFN- α_1 and IFN- α_2 are identical. The resulting ligation mixture was used to transform E.coli strain C600. Colonies were screened by Grunstein-Hogness hybridisation to the 500 bp Sau 3A - Sal I IFN- α_2 fragment (after nick translation). Plasmid DNA from eleven positive colonies (pIFS1202-1212) was isolated and digested with Eco RI and Sal I. All plasmids had single Eco RI and Sal I sites and treatment with both enzymes released an approximately 597bp DNA fragment, as expected. Plasmid pIFS1205 was sequenced as above.

Interferon expression

Detection of interferon production by E.coli cells containing plasmids with the synthetic IFN- α_2 gene was carried out initially using a standard antiviral assay [EMC virus on human amnion cells (WISH)] and an immunoradiometric assay (IRMA) with NK-2 monoclonal antibody. The results (table 1) show that the lac UV5/IFN- α_2 construction in E.coli strain PA340 grown under the conditions described produces a level of interferon activity just detectable by the IRMA. This level of interferon production could not be identified in the WISH/EMC antiviral assay. Four of the E.coli strain C600 colonies containing plasmids with the trp promoter/IFN- α_2 construction (pIFS1202-1205) were assayed (100ml cultures) for antiviral and NK2-IRMA activity. All four extracts had similar, substantial activity in both assays. In order to express the synthetic gene from the trp promoter under the most appropriate conditions, plasmid pIFS1205 was used to transform E.coli strain JA221. The resulting cultures were grown in tryptophan-limiting conditions to derepress the trp promoter. In view of the variations in interferon antiviral assays, a composite average result of a total of 6 assays from 3 separate 9L cultures is presented in Table 1. This we believe, gives a representative estimate of the level of interferon production in this

Table 1 IFN activity of *E.coli* extracts

Plasmid	<i>E. coli</i> strain	Promoter	Titre ^a	
			Antiviral	IRMA
p1101	PA 340 ^b	<u>lac</u> UV5	n.d.	2.5x10 ³
p1205	JA 221 ^c	<u>trp</u>	1.8x10 ⁷	5.4x10 ⁷

n.d. not detected

(a) Antiviral and IRMA units have been standardised to international units /L of culture at 1.0 OD units.

(b) Grown on L-broth

(c) Grown on minimal medium with added tryptophan (2µg/ml)

system.

The interferon protein expressed from plasmid pIFS1205 was characterised in three different systems. Extracts from lysates of transformed *E. coli* strain JA221 were chromatographed on the NK-2 Sepharose monoclonal antibody affinity support. All antiviral activity was retained, and the subsequently eluted interferon protein was compared with a sample of authentic IFN- α_2 by PAGE-SDS using Coomassie blue and silver stains (Fig 5). Lysates of ³⁵S-methionine-labelled preparations of minicells from *E.coli* strain DS410 transformed with plasmid pIFS1205 were run on the same gel but revealed by fluorography (Fig.5). The major interferon products from both systems had similar mobilities and corresponded with the authentic IFN- α_2 . Finally, an interferon protein of the same mobility was also produced when plasmid pIFS1205 was expressed in an *in vitro* transcription/translation system (A Charles and R Camble, data not shown). The IFN protein from minicells and the *in vitro* system were also both retained on NK-2 Sepharose and eluted in the same way as the product from JA221.

Further evidence of interferon production by *E.coli* cells containing the trp/IFN- α_2 plasmid was obtained by comparing the partially purified gene product (see Materials and Methods) with Namalwa interferon in a Daudi cell growth inhibition assay (36) (Fig. 6a) and a natural killer cell assay (2,36) (Fig. 6b). In both assays the synthetic IFN- α_2 gene product from *E.coli* displayed biological activity similar to that of Namalwa interferon. An extract of *E.coli* which did not contain plasmid bearing the synthetic gene, but which was processed in a similar manner to that for the IFN- α_2 gene product, was inactive in these assays.

Fig.5. Comparison of IFN- α_2 expressed from the synthetic gene in *E.coli* and minicells with authentic IFN- α_2 on an SDS-polyacrylamide gel. Lanes :1 and 7, [^{14}C]methylated protein markers; 2, lysate from [^{35}S]-methionine-labelled minicells containing plasmid pSTP1 (no interferon gene); 3, Lysate from [^{35}S]-methionine-labelled minicells containing plasmid pIFS1205; 4, authentic IFN- α_2 ; 5 and 6, IFN- α_2 from the synthetic gene after purification of *E.coli* lysate supernatants by immunoaffinity chromatography on NK-2-Sepharose.

After electrophoresis the gel was cut in two. The bands in lanes 1-3 were located by fluorography and the bands in lanes 4-7 were revealed by silver staining. The bands in the two pieces of gel were then realigned by lining up the two lanes of radiolabelled marker proteins, one revealed by fluorography and the other by silver stain.

DISCUSSION.

We have described here and elsewhere (11) the rapid chemical synthesis of large numbers of oligonucleotides and their stepwise ligation to furnish genes encoding relatively large proteins. Specific alterations to the sequences of such genes, either by resynthesis with selected different oligonucleotides or using point mutation techniques, should lead to modifications in the amino acid sequences of the encoded proteins. The biological properties of such variants may differ in some desirable way from those of the original molecule. The distinct antiviral activities of novel hybrid interferon molecules support this concept (8,9).

In the synthetic design of the IFN- α_2 gene we have adopted a flexible approach which should be applicable to the synthesis of any analogue gene by the introduction of modified sequences. The initial ligation of groups of oligonucleotides as shown in Fig. 3a produces double-stranded DNA with

Fig.6. (a) Dose response curve of the inhibition of DNA synthesis in Daudi cells and (b) dose-dependent enhancement of natural cytotoxicity of human peripheral blood lymphocytes by IFN- α_2 from the synthetic gene (solid circles) and Namalwa IFN (open circles). A mock extract of *E.coli*, adjusted to the same protein concentration as IFN- α_2 had no effect (triangles) in either assay.

component chains of unequal length. Analysis of the reaction mixtures by gel electrophoresis under denaturing conditions identifies the longer product strands unequivocally. The groups of 3+2 oligonucleotides (each with 15 nucleotide units encoding 5 amino acids) allow individual areas covering 10 or 15 amino acids to be independently manipulated. In preparative scale synthesis of these DNA sub-fragments the free 5' ends can be left unphosphorylated then labelled with [γ -³²P] ATP to a desired specific activity when required for use at a later stage.

Another feature of the synthetic design is that the codons used are chosen to be compatible with the host expression system. They might also, of course, be varied to determine whether this affects the amount or structure of the protein being expressed. It has been observed that when a high-abundance protein is to be produced in *E. coli* certain codons corresponding to rare tRNAs are avoided (37) since they may introduce translational errors (40). It is, therefore, probably desirable for a synthetic gene being expressed in *E. coli* to also avoid using these codons. Table 2 compares the codon usage in the synthetic IFN- α_2 gene with a natural IFN- α_2 gene produced by editing of cDNA (1). The sythentic gene contains no codons corresponding to minor tRNAs (asterisked) and a total of 23 codons are not

Table 2 Codon usage in synthetic (and natural) human IFN- α_2

	U	C	A	G	
U	Phe { 2 (4) 8 (6) Leu { 0 (1) 0 (4)	Ser { 2 (4) 6 (2) 0 (2) 0 (0)	Tyr { 0 (1) 5 (4) Ochre 1 - Amber -	Cys { 2 (3) 2 (1) Opal - (1) Trp 2 (2)	U C A G
C	Leu { 0 (1) 1 (6) 0 (1)* 20 (8)	Pro { 0 (3) 0 (2) 0 (0) 5 (0)	His { 2 (2) 1 (1) Gln { 6 (4) 6 (8)	Arg { 5 (0) 5 (0) 0 (0)* 0 (0)*	U C A G
A	Ile { 2 (1) 6 (6) 0 (1)* Met 5 (5)	Thr { 6 (3) 4 (4) 0 (3) 0 (0)	Asn { 0 (2) 4 (2) Lys { 8 (4) 2 (6)	Ser { 0 (2) 6 (4) Arg { 0 (6)* 0 (4)*	U C A G
G	Val { 3 (1) 1 (2) 3 (0) 0 (4)	Ala { 4 (4) 2 (2) 2 (2) 0 (0)	Asp { 1 (2) 7 (6) Glu { 12 (6) 2 (8)	Gly { 3 (1) 2 (1) 0 (1)* 0 (2)*	U C A G

used at all. There is, as yet, no evidence that this results in higher expression of IFN- α_2 or improved fidelity of amino acid sequence. The literature contains little amino acid sequence characterisation of the range of proteins which have been expressed from edited eukaryotic cDNAs in *E.coli*. Antigenicity of proteins produced by recombinant DNA techniques may arise if trace amounts of protein(s) differing in one or more amino acids are formed. Further studies are in progress to examine these problems.

Fig.7. Possible conformation for the mRNA produced from the *lac* UV5/IFN- α_2 construction.

The expression of eukaryotic genes in bacteria requires a promoter and translation initiation signal. These criteria are fulfilled with the lac UV5 construction. The poor expression observed here with this system (Table 1) might be explained in terms of mRNA secondary structure. Gheysen *et al* (41) have described variations in the nucleotide sequence between the Shine-Dalgarno sequence and a translation initiation codon for the SV-40 small-t antigen gene which lead to considerable differences in the production of antigen protein. Secondary-structure models for mRNAs suggest that an accessible start codon and ribosome interaction site are required for efficient translation initiation. Similar observations on the effects of mRNA secondary structure on translation initiation were reported by Hall *et al* (42).

A possible conformation for the mRNA produced from the lac UV5/IFN- α_2 construction is shown in Fig. 7. In this model, neither the Shine-Dalgarno sequence (AGGA) nor the initiation codon (AUG) is freely accessible. An identical structure is present in a lac UV5/IFN- α_1 construction described previously (39). In that example, however, translation initiation may occur to some extent at a position further downstream (P.A.Meacock, manuscript in preparation) resulting in a protein sharing many of the biological properties of natural IFN- α_1 .

In conclusion, we have described the total synthesis and expression of a human IFN- α_2 gene and some characterisation and biological properties of the resultant protein. The synthesis is designed to allow the facile generation of synthetic analogue IFN- α_2 genes for structure/activity analysis and incorporates codons thought to be efficiently expressed in bacteria. The gene is also amenable to *in vitro* mutagenesis approaches for the generation of specific or random modifications. Finally, given that the termini of the synthetic fragment can be changed at will during assembly, the possibility exists for further optimisation of features to the 3'- and/or 5'- sides of these genes (e.g. Shine-Dalgarno - initiation codon separation, promoter sequence, transcriptional terminators etc. etc.) with a view to maximising expression of IFN- α_2 or analogues thereof.

ACKNOWLEDGEMENTS

The authors wish to thank Professor C Weissmann for a sample of authentic IFN- α_2 , Dr D.Gilman for computer analysis, Mr C.G. Jones and Mr A.R.Pitt for technical assistance and Drs P.A.Meacock, J.Windass and D.Holland for many useful comments on this work.

REFERENCES

1. Streuli, M., Nagata, S. and Weissmann, C. (1980) *Science* 209, 1343-1347.
2. Lee, S.H., Kelley, S., Chiu, H. and Stebbing, N. (1982) *Cancer Res.* 42, 1312-1316.
3. Rehberg, E., Kelder, B., Hoal, E.G. and Pestka, S. (1982) *J.Biol.Chem.* 257, 11497-11502.
4. Taylor-Papadimitriou, J. and Balkwill, F.R. (1982) *Biochim.Biophys.Acta* 695, 49-67.
5. Goeddel, D.V., Leung, D.W., Dull, T.J., Gross, M., Lawn, R.M., McCandliss, R., Seeburg, P.M., Ullrich, A., Yelverton, E. and Gray, P.W. (1981) *Nature* 290, 20-26.
6. Brack, C., Nagata, S., Mantei, N and Weissmann, C. (1981) *Gene* 15, 379-394.
7. Weck, P.K., Apperson, S., May, L. and Stebbing, N. (1981) *J.gen.Virol.* 57, 233-237.
8. Streuli, M., Hall, A., Boll, W., Stewart, W.E. II, Nagata, S. and Weissmann, C. (1981) *Proc.Natl.Acad.Sci. U.S.A.* 78, 2848-2852.
9. Weck, P.K., Apperson, S., Stebbing, N., Gray, P.W., Leung, D., Shepard, H.M. and Goeddel, D.V. (1981) *Nucleic Acids Res.* 9, 6153-6166.
10. Here and throughout the paper, 'oligonucleotide' refers to oligodeoxyribonucleotide.
11. Edge, M.D., Greene, A.R., Heathcliffe, G.R., Meacock, P.A., Schuch, W., Scanlon, D.B., Atkinson, T.C., Newton, C.R. and Markham, A.F. (1981) *Nature* 292, 756-762.
12. Gait, M.J., Singh, M., Sheppard, R.C., Edge, M.D., Greene, A.R., Heathcliffe, G.R., Atkinson, T.C., Newton, C.R. and Markham, A.F. (1980) *Nucleic Acids Res.* 8, 1081-1096.
13. Markham, A.F., Edge, M.D., Atkinson, T.C., Greene, A.R., Heathcliffe, G.R., Newton, C.R. and Scanlon, D.B. (1980) *Nucleic Acids Res.* 8, 5193-5205.
14. Clarke, L. and Carbon, J. (1978) *J.Mol.Biol.* 120, 517-532.
15. Appleyard, R.K. (1954) *Genetics* 39, 440-452.
16. Reeve, J.N. (1977) *Molec. gen. Genet.* 158, 73-79.
17. Berberich, M.A. (1972) *Biochem.Biophys.Res.Comm.* 47, 1498-1503
18. Twigg, A.J. and Sherratt, D.J. (1980) *Nature* 283, 216-218.
19. Birnboim, H.C. and Doly, J. (1979) *Nucleic.Acids.Res.* 7, 1513-1523.
20. Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.* 114, 193-197.
21. Smith, H.O. (1980) *Meth.Enzymol.* 65, 371-380.
22. Maxam, A.M. and Gilbert, W. (1980) *Meth.Enzymol.* 65, 499-560.
23. Wallace, R.B., Johnson, M.J., Suggs, S.V., Miyoshi, K., Bhatt, R. and Itakura, K. (1981) *Gene* 16, 21-26.
24. Grunstein, M. and Hogness, D.S. (1975) *Proc.Natl.Acad.Sci. U.S.A.* 72, 3961-3965.
25. Windass, J.D. and Brammar, W.J. (1979) *Molec.gen.Genet.* 172, 329-337.
26. Miller, J.H. (1972) *Experiments in Molecular Genetics*, Cold Springs Harbor Laboratory, New York.
27. Newton, C.R., Greene, A.R., Heathcliffe, G.R., Atkinson, T.C., Holland, D., Markham, A.F. and Edge, M.D. (1983) *Anal.Biochem.* 129, 22-30.
28. Lennox, E.S. (1955) *Virology* 1, 190-206.
29. Spizizen, J. (1958) *Proc.Natl.Acad.Sci. U.S.A.* 44, 1072-1078.
30. Windass, J., Newton, C.R., De Maeyer-Guignard, J., Moore, V.E., Markham, A.F. and Edge, M.D. (1982) *Nucleic Acids Res.* 10, 6639-6657.
31. Camble, R., Petter, N.N. and Trueman, P. Unpublished data.
32. Secher, D.S. and Burke, D.C. (1980) *Nature* 285, 446-450.
33. Laemmli, U.K. (1970) *Nature* 227, 680-685.
34. Wray, W., Boulikas, T., Wray, V.P. and Hancock, R. (1981) *Anal.Biochem.* 118, 197-203.

35. Bonner, W.M. and Laskey, R.A. (1974) *Eur.J.Biochem.* 46, 83-88.
36. Masucci, M.G., Szigeti, R., Klein, E., Klein, G., Gruest, J., Montagnier, L., Taira, H., Hall, A., Nagata, S. and Weissmann, C. (1980) *Science* 209, 1431-1435.
37. Gouy, M. and Gautier, C. (1982) *Nucleic Acids Res.* 10, 7055-7074.
38. Frank, R. and Koster, H. (1979) *Nucleic Acids Res.* 6, 2069-2087.
39. De Maeyer, E., Skup, D., Prasad, K.S.N., De Maeyer-Guignard, J., Williams, B., Meacock, P., Sharpe, G., Pioli, D., Hennem, J., Schuch, W. and Atherton, K. (1982) *Proc.Natl.Acad.Sci. U.S.A.* 79, 4256-4259.
40. Grosjean, H. and Fiers, W. (1982) *Gene* 18, 199-209.
41. Gheysen, D., Iserentant, D., Derom, C. and Fiers, W. (1982) *Gene* 17, 55-63.
42. Hall, M.N., Gabay, J., Debarbouille, M. and Schwartz, M. (1982) *Nature* 295, 616-618.