The EMBO Journal Vol.2 No.12 pp.2349–2353, 1983

Cloning and structure of the human interleukin 2 chromosomal gene

Wim Degrave, Jan Tavernier, Fred Duerinck, Geert Plaetinck, Rene Devos[1] and Walter Fiers*

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

Communicated by W.Fiers
Received on 8 September 1983

Southern hybridization using ^{32}P-labelled human interleukin 2 (IL2) cDNA probes revealed the existence of a single human IL2 gene. Five clones containing the human IL2 chromosomal gene were isolated from two different human DNA libraries cloned in either λ Charon 4A or L47 phages. Analysis of the clones showed that they contained different, overlapping portions of human DNA which were derived from the same chromosomal segment. Restriction fragments which hybridized with labelled IL2 cDNA probes were subcloned into plasmid pUR250 and the sequence and organization of the IL2 gene was determined. It contains three introns, 90 bp, ±2400 bp and ±1900 bp in length, respectively. The organization of the genomic clone resembles that of another lymphokine, interferon-γ, but no clear homology was found by comparing either the coding sequence or the 5′- and 3′-flanking sequences of the two genes.
Key words: genomic sequence/human interleukin 2/lymphokine/polymerase II promoter

Introduction

Human interleukin 2 (IL2, also known as TCGF or T-cell growth factor) is synthesized and secreted by activated T-lymphocytes and plays an important immunoregulatory role. It stimulates DNA synthesis and the long-term *in vitro* proliferation of T-cells following interaction with antigen in the proper context (Morgan *et al.*, 1976; Mier and Gallo, 1980; Smith, 1980). IL2 also seems to be involved in the activation of natural killer (NK) cells (Dennert, 1980; Henney *et al.*, 1981). The natural product is probably a glycoprotein with a mol. wt. of ~15 000 daltons. Recently, we reported the cloning, characterization and expression of the human IL2 cDNA gene (Devos *et al.*, 1983). The unique coding region of the preprotein is 153 amino acids long and its 20 N-terminal residues are presumed to represent a signal peptide. The deduced amino acid sequence was identical to that previously reported by Taniguchi *et al.* (1983). IL2 was expressed efficiently in *Escherichia coli* under the control either of the tryptophan promoter or the phage λ P_L promoter.

Here we report the isolation and characterization of the human IL2 chromosomal gene. This gene was obtained from two human gene libraries, cloned in λ Charon 4A and in L47 phages. The data obtained reveal the presence of three introns in the IL2 chromosomal gene but little, if any, sequence homology to another previously characterized human lymphokine gene, interferon-γ (IFN-γ) (Taya *et al.*, 1982).

[1]Present address: Biogent, Plateaustraat 22, B-9000 Ghent, Belgium.
*To whom reprint requests should be sent.

Results

Southern blotting of digested, total human genomic DNA

High mol. wt. human genomic DNA was digested with four different restriction enzymes (*Eco*RI, *Hind*III, *Pvu*II and *Pst*I), which do not cut within the IL2 cDNA (Devos *et al.*, 1983). After electrophoresis on a 1% agarose gel and transfer to nitrocellulose, DNA was hybridized with a denatured ^{32}P-labelled IL2 cDNA-specific *Hinf* fragment covering a large part of the coding sequence and of the 3′-untranslated region (Devos *et al.*, 1983). One single band was detected in the case of *Eco*RI, *Pvu*II and *Pst*I digestion with a length of ~3800 bp, 9000 bp and 12 000 bp, respectively, while the *Hind*III-digested DNA gave two bands (data not shown). The results are consistent with the presence of a single chromosomal IL2 gene, containing at least one intron. Also, hybridization under low stringency conditions did not reveal additional related genes.

Isolation of an IL2 gene from a human genome library

Two human gene banks, prepared according to Lawn *et al.* (1978) and to Loenen and Brammar (1980) respectively, were plated out on *E. coli* BHB 2600 and screened *in situ* (Benton and Davis, 1977) using a ^{32}P-labelled IL2 cDNA-specific *Hinf* fragment (see Devos *et al.*, 1983). Two λ recombinants (λCH4A-gHil2-1 and -2) were isolated after screening 1 x 10^6 recombinant phages from the first gene bank and three recombinants (L47-gHil2-1, -2 and -3) were found after screening 7 x 10^5 phages of the second gene bank.

Analysis of the recombinant phages and subcloning

Digestion of the phage DNAs with a series of restriction enzymes, Southern blotting and hybridization with an IL2 cDNA-specific *Hinf* fragment or with a small *Hinf-Hae*III subfragment, indicated that all the phage inserts were derived from the same chromosomal segment, although they contained shifted portions of genomic DNA.

One of the phages (λCH4A-gHil2-1) was shown to contain the total human IL2 gene as two *Eco*RI restriction fragments of 3800 bp and 3000 bp long. These two *Eco*RI fragments were subcloned in the plasmid pUR-250 (Rüther, 1982). Obviously, the 3000-bp *Eco*RI fragment must have originated by addition of an *Eco*RI linker on an internal *Alu*I or *Hae*III restriction site, cleaved during the construction of the recombinant phage, since the Southern blotting of the *Eco*RI-digested, total human genomic DNA revealed only a single band which undoubtedly consisted of two fragments, ~3800 bp long.

Restriction analysis and direction of transcription

To elucidate the structure of the IL2 gene, we analyzed in further detail the cross-hybridizing *Eco*RI 3000-bp and 3800-bp fragments of λCH4A-gHil2-1, which gave rise to the subclones pUR-gHil2-1 and -2, respectively.

^{32}P-Labelled probes were prepared by nick-translation of two DNA fragments isolated from the IL2 cDNA clone pSV-Hil2-0 (see Devos *et al.*, 1983); namely, on the one hand, a *Hgi*AI-*Xba*I fragment (~175 bp long) which contains the 5′ half of the IL2 coding region, and on the other hand an

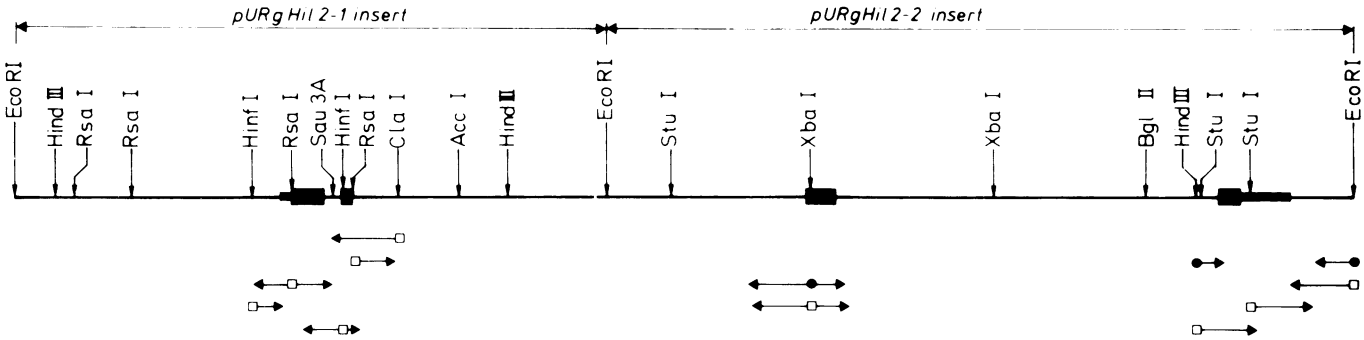

Fig. 1. Restriction map and sequencing strategy for the human IL2 chromosomal gene. Black boxes show the coding regions while thick bars represent 5′-and 3′-untranslated regions. Above, restriction sites of the plasmid inserts, which were used for sequencing, are shown. *Rsa*I, *Hinf*I and *Sau*3A maps are not complete. The sequence strategy is given below. Open squares represent 5′ end-labelling and closed circles refer to 3′ end-labelling. (We have not proven that the two *Eco*RI fragments are adjacent.)

*Xba*I-*Stu*I fragment (~260 bp long) which spans the 3′-terminal half of the coding region as well as a small part of the 3′-non-coding information. These probes were used to analyze the plasmids pUR-gHil2-1 and -2 by Southern blotting, after digestion with several restriction enzymes.

The *Eco*RI 3000-bp insert (pUR-gHil2-1) hybridized only with the *Hgi*AI-*Xba*I probe while the *Eco*RI 3800-bp insert (pUR-gHil2-2) hybridized with both probes. More refined restriction mapping followed by Southern blot analysis resulted in tentative restriction maps of both plasmid inserts and indicated the direction of transcription as shown in Figure 1.

Nucleotide sequence of the IL2 gene

Figure 1 also shows the strategy followed for DNA sequencing. The primary structure of the genomic IL2 gene and of the flanking regions is shown in Figure 2. Plasmid pUR-gHil2-1, indeed, contained the promoter region and the beginning of the IL2 gene. As reported previously, (Devos *et al.*, 1983), the adenyl residue at nucleotide 1 presumably constitutes the mRNA initiation site. A TATAAT consensus sequence is present at 32 nucleotides upstream from this putative first nucleotide of the IL2 mRNA. The first intron is only 90 bp long and starts after nucleotide 194 of the cDNA sequence. The second exon, 60 nucleotides long, is followed by the second intron, which starts after nucleotide 254 of the IL2 mRNA. The 3′ acceptor splice site of the second intron was located in the plasmid pUR-gHil2-2, followed by the third exon which was 144 bp long. The third intron starts after nucleotide 398 and is ~1900 bp long. It is followed by the remaining part of the coding information, the uninterrupted 3′ region and the sequence downstream from the polyadenylation site. Each intron interrupts the reading frame precisely between codons. The nucleotide sequence which we have determined for the IL2 cDNA (Devos *et al.*, 1983) is in complete agreement with the sequence data obtained from the genomic clone, except for the nucleotide G at position 161 in the IL2 cDNA (Devos *et al.*, 1983; Taniguchi *et al.*, 1983) which we determined to be T in our genomic clone. This substitution does not lead to a change in the amino acid sequence and is presumably due to allelic variation.

Discussion

The cloning and characterization of the IL2-cDNA gene has previously been reported (Taniguchi *et al.*, 1983; Devos *et al.*, 1983) as was the IFN-γ cDNA gene (Devos *et al.*, 1982; Gray *et al.*, 1982) and the IFN-γ chromosomal gene (Gray and Goeddel, 1982; Taya *et al.*, 1982). It was shown that IFN-γ was present as a single gene in the human genome and that it contained three introns (Figure 3). As described above, Southern blotting of total, digested human genomic DNA also suggested the existence of one single IL2 gene. Further detailed restriction site mapping and sequence analysis revealed a single gene containing three introns, 90 bp, ±2400 bp and ±1900 bp long, respectively. The IL2 chromosomal gene organization is shown in Figure 3 and is similar to that of IFN-γ. Although IFN-γ and IL2 are induced concomitantly in mitogen-treated lymphocyte cultures and both are secreted by T-lymphocytes, no significant homology was found in their cDNA sequence nor in the 5′- and 3′-flanking regions.

At most, two regions with a moderate homology were found in the promoter region (Figure 4). The first segment from nucleotide −20 to −40 comprises the TATAAAT box (reviewed by Breathnach and Chambon, 1981). The second segment, from nucleotide −60 to −85, contains a sequence GTAAAACATT which somewhat resembles the consensus sequence GG$^{C}_{T}$CAATCT (Benoist *et al.*, 1980), present around position −70 to −80 in several viral and cellular promoter regions.

Both in the IL2 and in the IFN-γ gene, each intron interrupts the reading frame precisely between codons. Furthermore, there is a clear similarity in their donor and acceptor splice sites (Figure 5). In each case the GT/AG splice rule (Mount, 1982) is followed. Remarkably, both for IL2 and IFN-γ, the donor and acceptor sequence of the first intron is a rather rare exception in that there is no direct repeat present in the donor *versus* acceptor splice site (hence there is no ambiguity as to where exactly the junction occurs), while for the splice junctions of the second and third intron, the repeat length is three nucleotides for both IL2 and IFN-γ.

The fact that both IL2 and IFN-γ are lymphokines produced by stimulated T-cells, both are growth regulatory factors, both have approximately the same size (on denaturing gels), and since they have similar genomic organization, this possibly could mean that they are derived from a common ancestral gene. No convincing homology can be deduced from the present primary nucleotide or amino acid sequence, but perhaps the elucidation of the three-dimensional structures of these remarkable proteins may shed some light on this question.

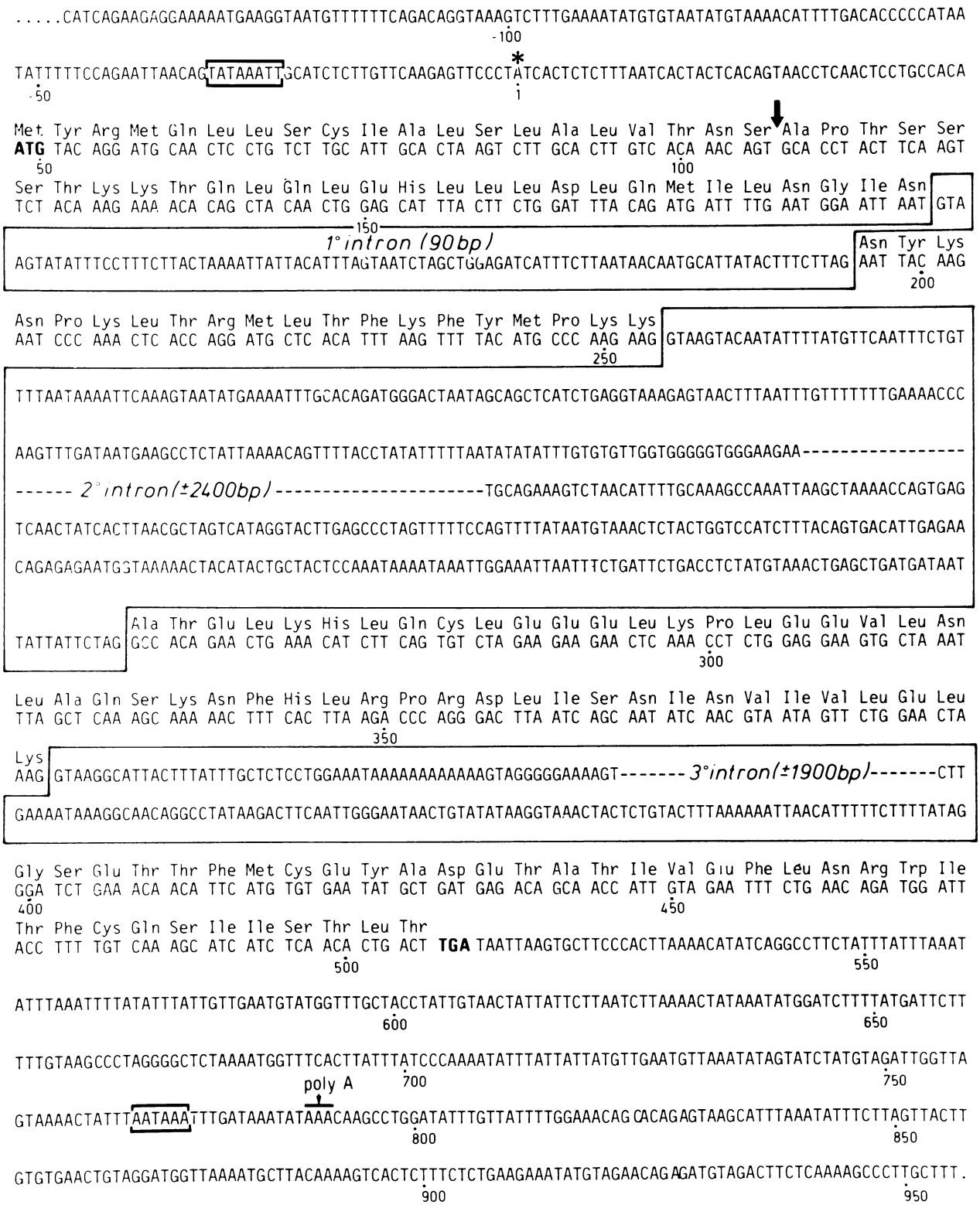

Fig. 2. Nucleotide sequence of the human IL2 chromosomal gene and of the flanking regions. The sequence, excluding the introns, was numbered starting from the putative mRNA start site (indicated by an asterisk). The arrow indicates the presumed signal sequence cleavage site (Devos *et al.*, 1983). A TATAAATT- sequence, present 32 nucleotides upstream from the putative start site, as well as the poly(A) addition signal AATAAA, are indicated by heavy bars. The three introns (only the sequence of the first small intron is given in total) are boxed.

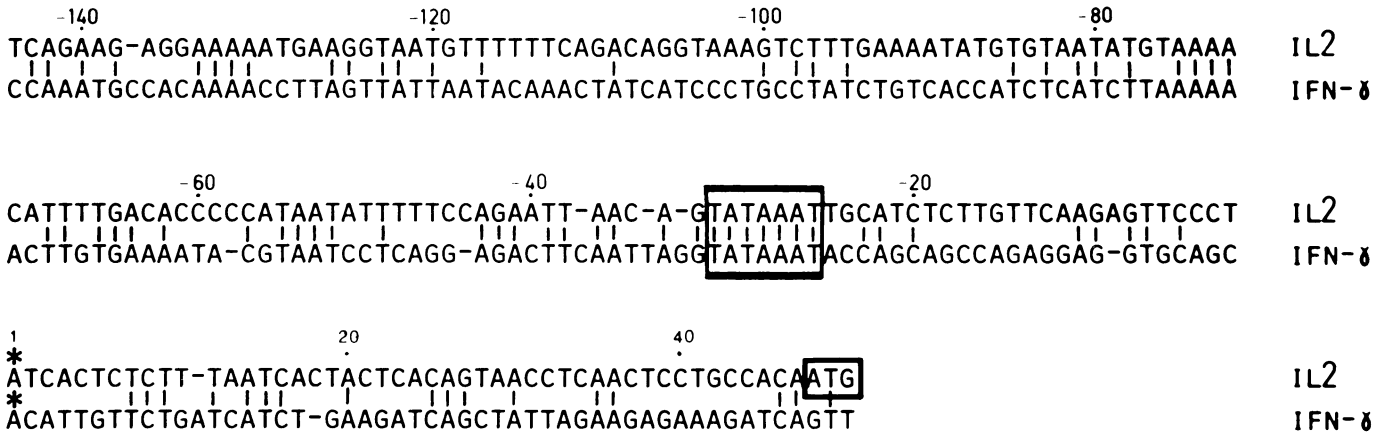

Fig. 3. Schematic representation of the organization of the chromosomal IL2 and IFN-γ genes. Black boxes show the coding region. The 5′- and 3′-untranslated regions of the genes are shown as hatched boxes. Important transcription and translation signals are indicated below. The numbers refer to the length of the segments (in base pairs).

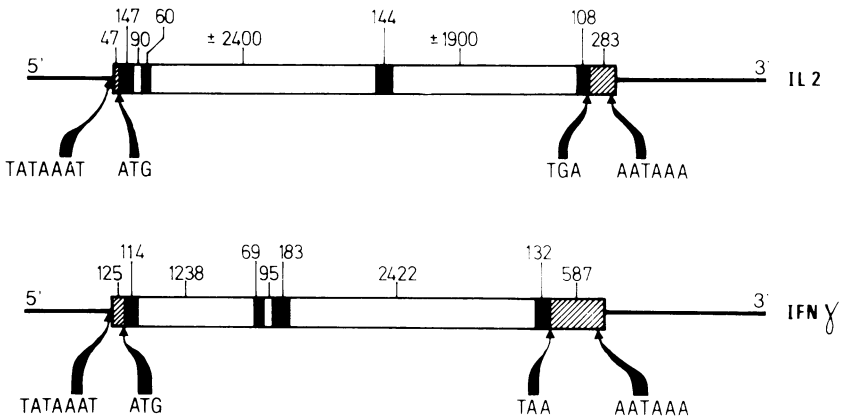

Fig. 4. Comparison of the 5′-flanking region including the putative promoter of the human IL2 and IFN-γ genes. Nucleotides, identical in both sequences are connected by vertical bars. Gaps are created to show maximal homology. The numbering for the IL2 gene is as shown in Figure 2.

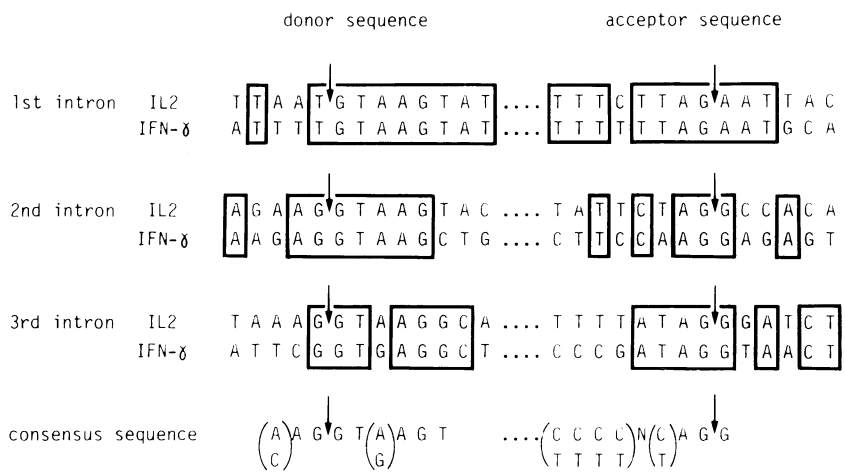

Fig. 5. Comparison of the splice junctions of the IL2 and IFN-γ genes and the consensus sequence. The arrows indicate the splice sites assuming that the GT/AG rule is followed. Homologous nucleotides between the two genes are boxed.

Materials and methods

Procedures for preparing DNA

High mol. wt. DNA from human spleens and plasmid DNA were prepared as described in Tavernier *et al.* (1981). Phage DNA was prepared essentially as described by Maniatis *et al.* (1982).

Phage library screening

A λ Charon 4A phage recombinant library (*Alu*I and *Hae*III partial digestion fragments), constructed by Lawn *et al.* (1978), and a library (*Mbo*I partial digestion fragments) cloned in a λL47 vector according to Loenen and Brammar (1980) were used. Approximately 1 000 000 and 750 000 phages, respectively, were screened as described by Tavernier *et al.* (1981), using an

IL2 cDNA-specific *Hinf* fragment (450 bp) derived from pAT153-Hil2 (Devos *et al.*, 1983), which was ^{32}P-labelled by nick-translation.

Two and three plaques, respectively, were picked out and were plaque purified. The recombinant phages were grown in 500 ml cultures for preparation of DNA.

Restriction endonuclease digestion, agarose gel electrophoresis, Southern blotting and hybridization were carried out as described before (Tavernier *et al.*, 1981).

Subcloning, restriction mapping and DNA sequencing

*Eco*RI fragments were subcloned into a single *Eco*RI site in the β-galactosidase α-peptide gene region of the plasmid pUR-250 (Rüther, 1982). *E. coli* RRIZΔM15 was used for transformation.

Restriction maps with *Rsa*I and *Hinf* were made by the partial digestion procedure of Smith and Birnstiel (1976).

DNA sequencing was done using the chemical method of Maxam and Gilbert (1980).

Acknowledgements

We thank Dr.T.Maniatis and Dr.M.Ballivet for kindly providing the human DNA libraries in the vectors λ Charon 4A and L47, respectively. We are also grateful to Rose Wulgaert for plasmid preparation and Fred Shapiro and Wim Drijvers for their help in preparing this manuscript. G.P. holds a fellowship from the IWONL. This research was supported by a grant from the Fonds voor Geneeskundig Wetenschappelijk Onderzoek and from Biogen S.A.

References

Benoist,C., O'Hare,K., Breathnach,R. and Chambon,P. (1980) *Nucleic Acids Res.*, **8**, 127-142.

Benton,W.D. and Davis,R.W. (1977) *Science (Wash.)*, **196**, 180-182.

Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349-383.

Dennert,G. (1980) *Nature*, **287**, 47-49.

Devos,R., Cheroutre,H., Taya,Y., Degrave,W., Van Heuverswyn,H. and Fiers,W. (1982) *Nucleic Acids Res.*, **10**, 2487-2501.

Devos,R., Plaetinck,G., Cheroutre,H., Simons,G., Degrave,W., Tavernier,J. Remaut,E. and Fiers,W. (1983) *Nucleic Acids Res.*, **11**, 4307-4323.

Gray,P.W. and Goeddell,D.V. (1982) *Nature*, **298**, 859-863.

Gray,P.W., Leung,D.W., Pennica,D., Yelverton,E., Najarian,R., Simonsen, C.C., Derynck,R., Sherwood,P.J., Wallace,D.M., Berger,S.L., Levinson, A.D. and Goeddell,D.V. (1982) *Nature*, **295**, 503-508.

Henney,C.S., Kuribayashi,K., Kern,D.E. and Gillis,S. (1981) *Nature*, **291**, 335-338.

Lawn,R.M., Fritsch,E.F., Parker,R.C., Blake,G. and Maniatis,T. (1978) *Cell*, **15**, 1157-1174.

Loenen,W.A.M. and Brammar,W.J. (1980) *Gene*, **10**, 249-259.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY, pp. 76-85.

Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.

Mier,J.W. and Gallo,R.C. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 6134-6138.

Morgan,D.A., Ruscetti,F.W. and Gallo,R.C. (1976) *Science (Wash.)*, **193**, 1007-1008.

Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459-472.

Rüther,U. (1982) *Nucleic Acids Res.*, **10**, 5765-5772.

Smith,H.O. and Birnstiel,M.L. (1976) *Nucleic Acids Res.*, **3**, 2387-2398.

Smith,K.A. (1980) *Immunol. Rev.*, **51**, 337-357.

Taniguchi,T., Matsui,H., Fujita,T., Takaoka,C., Kashima,N., Yoshimoto, R. and Hamuro,J. (1983) *Nature*, **302**, 305-310.

Tavernier,J., Derynck,R. and Fiers,W. (1981) *Nucleic Acids Res.*, **9**, 461-471.

Taya,Y., Devos,R., Tavernier,J., Cheroutre,H., Engler,G. and Fiers,W. (1982) *EMBO J.*, **1**, 953-958.