Structure of the human interleukin 2 gene

(Southern blotting/genomic clone/gene structure/sequence homology/lymphokine gene expression)

TAKASHI FUJITA, CHIKAKO TAKAOKA, HIROSHI MATSUI*, AND TADATSUGU TANIGUCHI

Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Toshima-ku, Tokyo 170, Japan

Communicated by Werner Henle, September 8, 1983

ABSTRACT We have cloned two species of *EcoRI*-cleaved DNA segments that together cover the entire sequence for the human interleukin 2 gene and have determined the nucleotide sequence of the gene and its flanking regions. The gene contains three introns and the exon sequences can be aligned with the previously reported cDNA sequence almost perfectly except for a few nucleotides in the 3' nontranslated region. The promoter region contains a prototype "TATA" sequence as well as a notable palindromic sequence. Particularly interesting is the presence of sequences in this region that are homologous to the promoter region of the human interferon- γ gene. In addition, a sequence that closely resembles the core sequence for the viral enhancer elements has been found in the second intron. Such sequences may play a role in the expression of the interleukin 2 gene in lectin- or antigen-stimulated T lymphocytes.

Interleukin 2 (IL-2) is a lymphokine produced by lectin- or antigen-activated T cells (1, 2). Among the various biological activities ascribed to IL-2 (1, 3-5), the principal role of this immunoregulatory molecule appears to be the stimulation of the proliferative response of activated T-cell clones and hence it plays a key role in the regulation of the T-cell clonal expansion (5).

We recently reported the isolation and nucleotide sequence of a cloned cDNA for human IL-2 (6). The deduced amino acid sequence of mature IL-2 showed very little homology with any humoral factors whose structures have been elucidated. Although human IL-2 is reportedly heterogeneous (7), we have been able to identify only a single species of mRNA coding for IL-2 in both a leukemia T-cell line (Jurkat-111) and normal peripheral blood lymphocytes (ref. 6; unpublished results).

Despite extensive studies on the biological characteristics of IL-2, little is known about the regulation of IL-2 expression. It is well documented that the production of IL-2 is specific for mature thymus-derived T cells whereas IL-2-responsive cells belong to a distinct subset (5). Since neither IL-2 activity nor IL-2-specific mRNA is produced by such producer cells without induction by lectins or antigens (1, 6, 8), IL-2 expression is likely to be controlled at the mRNA transcription level.

To study the structural organization and to obtain further information about the expression of the human *IL-2* gene, we have cloned and analyzed the *IL-2* chromosomal gene and its flanking regions.

MATERIALS AND METHODS

Preparation and Blotting Analysis of Human DNA. Human chromosomal DNA was extracted as described (9) from the peripheral blood lymphocytes of healthy donors (10) or from Jurkat-111 cells (6). Blotting analysis of the total DNA was carried out according to the procedure of Southern (11).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Cloning and Sequence Analysis of the *IL-2* Gene. DNA (500 μ g) from peripheral blood lymphocytes was digested with *EcoRI* and fractionated on a 10-40% linear sucrose gradient in 20 mM Tris-HCl, pH 8.0/1 M NaCl/5 mM EDTA by centrifugation at 27,000 rpm for 24 hr at 4°C in a Beckman SW28 rotor. DNA from the fractions of the $\approx$ 3.5-kilobase (kb) fragments containing the *IL-2* gene sequence (1.2 μ g) was ligated to 2.0 μ g of the purified arm of λ gt wes λ B phage, and the resulting recombinant phage DNA was packaged *in vitro* as described by Blattner *et al.* (12). IL-2-specific phage clones were screened by the *in situ* procedure (13) using the pIL2-50A cDNA insert (6) as the probe, and clones designated λ IL2Xba and λ IL2Taq were identified. For the purpose of sequence analysis, each cloned *EcoRI* fragment was excised from the phage DNA and introduced into pBR322 to yield p λ IL2Xba and p λ IL2Taq.

The nucleotide sequence of the cloned DNA was determined by the procedure of Maxam and Gilbert (14).

RESULTS

Blotting Analysis of the Chromosomal DNA. To study the organization and structure of the human *IL-2* gene, we first carried out Southern blot analyses of the chromosomal DNA from the normal peripheral blood lymphocytes and Jurkat-111 cells from which we had extracted mRNA for the *IL-2* cDNA cloning (6). As shown in Fig. 1, a rather simple pattern was obtained when nick-translated whole pIL2-50A cDNA was used as the probe. However, digestion of the DNA with *Bgl* II and *Hind* III, which do not cleave the *IL-2* cDNA (6), gave rise to two positive bands (13.0 and 7.2 kb for *Bgl* II, 3.4 and 2.2 kb for *Hind* III) and digestion with *Xba* I, which cuts the cDNA only once, gave three positive bands (6.8, 3.5, and 1.3 kb). On the other hand, *EcoRI* digestion yielded a single positive band (3.5 kb) by this analysis. No differences were observed among the tested DNAs and an identical pattern was obtained with the DNA of non-lymphoid cells (result not shown). When an identical filter was hybridized with the 5' portion of the cDNA (*Pst* I-*Xba* I fragment; ref. 6) as the probe, only a single band was positive with the *Bgl* II- or *Hind* III-digested DNA (7.2 kb for *Bgl* II and 2.2 kb for *Hind* III; results not shown). These observations thus indicated the presence of a single copy *IL-2* gene containing an intron(s).

Isolation of the Recombinant Phage Containing the *IL-2* Gene Sequence. Because *EcoRI* digestion of the total DNA gave a single positive band (3.5 kb) in the blotting analysis, we have purified this 3.5 kb DNA fragment by a sucrose gradient centrifugation (see *Materials and Methods*) and cloned into λ gt wes λ B phage. About 8×10^4 plaques were screened by the *in situ* procedure (13) using the 32 P-labeled pIL2-50A cDNA in-

Abbreviations: IL-2, interleukin 2; kb, kilobase(s); bp, base pair(s); IFN- γ , interferon- γ .

* Present address: Central Research Laboratories, Ajinomoto Co., Inc., Suzuki-cho, Kawasaki-ku, Kawasaki 210, Japan.

FIG. 1. Blot hybridization analysis of human chromosomal DNA. Chromosomal DNA was prepared from peripheral blood lymphocytes of two healthy donors (lanes a and b) or from Jurkat-111 cells (lanes c) (6). Each DNA sample (10 μ g) was digested with various restriction endonucleases, and the digests were electrophoresed on 0.8% agarose gels and blotted to a nitrocellulose filter (11). The 32 P-labeled DNA probe was prepared from the whole cDNA insert of pIL2-50A as described (6, 15); 1×10^7 cpm (specific activity, 10^8 cpm/ μ g) was hybridized with the filter for 16 hr and then washed as described (15). Numbers on the right indicate the positions of size markers. kbp, kilobase pairs.

sert as the probe, and six independent phage clones were identified. Restriction endonuclease analysis of these clones revealed the presence of two distinct groups of recombinant phages. Two recombinant phages, each containing an identical 3.5-kb insert that was cleaved twice by *Xba* I, and four phages, each containing an identical 3.5-kb insert without any *Xba* I cleavage sites, were obtained. The inserts were separately cloned in pBR322, and the resulting plasmids were designated pIL2Xba and pIL2Taq. Restriction endonuclease cleavage maps of the cloned inserts are presented in Fig. 2.

Nucleotide Sequence Analysis of the *IL-2* Gene. The nucleotide sequences of the whole 3,662-base-pair (bp) insert from pIL2Xba and a portion of the insert (2,102 bp) from pIL2Taq were determined by following the strategy shown in Fig. 2. The *IL-2* cDNA sequence previously determined (6) was aligned with the chromosomal gene sequence as shown in Fig. 3. The two sequences matched well except for a nucleotide (position

3,085 in Fig. 3b, position 503 of ref. 6) in the protein-encoding region. A nucleotide identical to that found in this position of the chromosomal gene has been found in three independent cDNA clones (unpublished results). In the 3' noncoding region sequence, A-A at position 665-666 of the pIL2-50A cDNA is not present and A-T at position 741-742 of the cDNA is replaced by guanosine (position 3,321 in Fig. 3b) in the cloned gene. The sequence analysis revealed that all four exons are in the two *Eco*RI fragments and that the gene contains three introns. Each intron interrupts the reading frame precisely between codons and contains the G-T and A-G consensus sequences (16, 17) at the 5' and the 3' termini, respectively.

The nucleotide sequence analysis of the *IL-2* gene also allows us to compare sequences thought to play a role in the gene expression. The first adenosine residue of the pIL2-50A cDNA, a presumed cap site, is aligned in the chromosomal gene at the position 32 bp downstream of the well-conserved "TATA" box, which determines the specificity of the mRNA synthesis initiation by RNA polymerase II (18). Other sequence elements of interest have also been found in the cloned DNA fragments (see *Discussion*).

Contiguity of the Cloned *Eco*RI Fragments in the Human Genome. Because the two *Eco*RI fragments were cloned independently from the total DNA, it is possible that the two fragments are interrupted by one or more *Eco*RI segments that would constitute a part of the intron. To test this, total human lymphocyte DNA was digested with *Hind*III and other restriction endonucleases as indicated in Fig. 4a and subjected to the blotting analysis by using a unique *Hind*III-*Eco*RI fragment from pIL2Taq (positions 1,566-2,096 of Fig. 3a) as the radioactive probe. As shown in Fig. 4b, unique hybridization patterns were obtained (the sizes of the positive fragments are calculated to be about 3,400 bp for *Hind*III, 1,600 bp for *Hind*III/*Xba* I, 1,300 bp for *Hind*III/*Hinc*II, and 850 bp for *Hind*III/*Hae* III, whereas the expected sizes of these fragments would be 3,407, 1,549, 1,314, and 808 bp, respectively). Although we cannot strictly rule out the possibility of the presence of a very small *Eco*RI fragment(s) between the DNA fragments, the re-

FIG. 2. Restriction enzyme cleavage map and sequence analysis strategy of the two *Eco*RI DNA fragments containing the *IL-2* gene sequence. The fragments were first cloned in λ gt10 phage and then transferred to pBR322. The restriction map was constructed as described (15). Cleavage sites for *Bst*NI in the pIL2Taq insert and for *Hinf*I in pIL2Xba are not indicated. Arrows indicate the direction and extent of sequence analysis of each fragment; the vertical bar at the end of each arrow represents 32 P-labeled 5'-terminal phosphate.

Fig. 3. Nucleotide sequences of the cloned *EcoRI* fragments containing the *IL-2* gene sequence. (a) pAIL2Taq. (b) pAIL2Xba. The sequence was determined according to the procedure of Maxam and Gilbert (14). The amino acid sequence of IL-2 polypeptide was deduced by aligning the nucleotide sequence with that of pIL2-50A cDNA (6). The asterisk (position 431 of a) indicates the presumed mRNA start site (6). The TATA box and A-A-T-A-A sequence are underlined. Arrows indicate the position of poly(A) attachment. The sequence homologous to the core sequence of the enhancer elements (see Discussion) is marked by a dashed line.

FIG. 4. Southern blot analysis of total human DNA with a 0.5-kb *Hind*III/*Eco*RI fragment from pIL2Taq as the hybridization probe. (a) Blotting strategy and expected results: If two *Eco*RI fragments are contiguous to each other, DNA fragments of expected sizes as shown here would become positive to *Hind*III/*Eco*RI fragment. (b) Results obtained: Total DNA from human lymphocytes was digested with *Hind*III, *Hind*III/*Xba* I, *Hind*III/*Hinc*II, and *Hind*III/*Hae* III (lanes 1–4) and samples were analyzed by the blotting procedure as described (15). Numbers on the right indicate the sizes of the respective positive DNA fragments (Fig. 3a) calculated from several size markers.

sult is best explained by assuming that two DNA fragments are contiguous with each other in the expected orientation (Fig. 4a).

DISCUSSION

We have analyzed the organization and structure of the human *IL-2* chromosomal gene. Combining the results of Southern blot analysis of the total human DNA (Fig. 1) and structural analysis

of the cloned *Eco*RI fragments (Figs. 2 and 3), we conclude that the *IL-2* gene exists in a single copy per haploid human genome. Thus, molecular heterogeneity of human *IL-2* (5, 7) is probably due to post-transcriptional modifications. So far, we have been unable to detect any sequence that cross-hybridizes to the *IL-2* cDNA probe in the human genome under lower stringent conditions for hybridization (results not shown).

Extensive nucleotide sequence analysis of the two *Eco*RI DNA fragments revealed that the *IL-2* gene is interrupted by three introns. By aligning the *IL-2* cDNA sequence with the chromosomal gene sequence, we found two regions of nucleotide divergence in the 3' noncoding region. Such divergences may represent genetic polymorphism in this region of the gene.

Most of the inducible genes of eukaryotes contain unique DNA sequences that mediate inducer-responsive gene expression upstream from the TATA box (18–23). From this point of view, it was of interest to compare the sequences in the 5' region of the *IL-2* gene with that of the interferon- γ (*IFN*- γ) gene, whose expression is under similar control. Like other lymphokines, production of *IFN*- γ is specific to T cells stimulated by mitogens or by antigens (24, 25). Several reports also suggest that induction of *IL-2*, *IFN*- γ , and interleukin-1 (*IL-1*) involve a cascade of interactions among them (26–28). As shown in Fig. 5, two blocks of sequence homologous to the 5' flanking region of the *IFN*- γ gene (29, 30) are present in the *IL-2* gene. It may also be worth noting that a sequence within one of the blocks ($^{-69}$ T-T-T-G-G-A-C-A-C-C-C-C-A-T-A-A $^{-51}$) potentially constitutes a stem-loop structure with its upstream sequence [T-T-T-G-G-G-G-G-T-T-T-A-A-A-G, position -175 to -161 from the cap site (position 257–271 of Fig. 3a)].

Recently, evidence for a novel function for an intron sequence in immunoglobulin heavy chain genes has been provided (31, 32). The sequence appears to function in a manner identical to viral enhancer elements (32–35) except that this effect occurs only in particular cell types (e.g., myeloma cells). The sequence also seems to show a certain degree of homology to viral enhancer elements for which a core sequence, G-G-T-G-T-G-G-A-A-A-G, has been proposed (34, 36). In view of these findings, it may be worth noting that a sequence, G-G-T-G-T-G-T-A-A-A-G, that is almost identical to such a core sequence is located within the second intron of the *IL-2* gene (position 1,636–1,646 of Fig. 3a).

The mechanism of lymphokine gene expression is totally unknown at present and it remains to be seen whether or not the above-mentioned sequences in fact play a role in the expression of the *IL-2* gene in T lymphocytes.

We thank Drs. M. Yoshida and J. Hamuro for providing us with the purified Agt w ϕ AB phage arm and Jurkat-111 cells, respectively, as

FIG. 5. Nucleotide sequence homology in the 5' flanking regions of the human *IL-2* and *IFN*- γ genes. In aligning the two sequences, gaps were introduced to maximize the homology. Numbers under the sequence refer to the nucleotide position from the mRNA start site (+1) of each gene. The TATA box is underlined. Sequence data for the *IFN*- γ gene are from Gray and Goeddel (29) and Taya *et al.* (30).

well as for valuable advice. We thank Dr. H. Sugano for his continuous support. Thanks are also due to Mr. T. Imada and Ms. S. Ichikawa for help in searching homologous sequences and preparing the manuscript, respectively. We are indebted to Dr. J. Vilček for valuable discussions and comments. This work was supported in part by Princess Takamatsu Cancer Research Funds.

1. Morgan, D. A., Ruscetti, F. W. & Gallo, R. (1976) *Science* **193**, 1003–1027.
2. Gillis, S., Felm, M. M., Ou, W. & Smith, K. J. (1978) *Immunology* **120**, 2023–2027.
3. Chen, B. M. & DiSabato, G. (1967) *Cell. Immunol.* **22**, 211–224.
4. Wagner, H., Hardt, C., Herg, K., Rollinghoff, M. & Pfizenmaier, K. (1980) *Nature (London)* **284**, 278–280.
5. Smith, K. A. (1980) *Immunol. Rev.* **51**, 337–357.
6. Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R. & Hamuro, J. (1983) *Nature (London)* **302**, 305–310.
7. Robb, R. J. & Smith, K. A. (1981) *Mol. Immunol.* **18**, 1087–1094.
8. Gillis, S., Baker, P. E., Ruscetti, F. W. & Smith, K. A. (1978) *J. Exp. Med.* **148**, 1093–1100.
9. Ohno, S. & Taniguchi, T. (1982) *Nucleic Acids Res.* **10**, 967–977.
10. Taniguchi, T., Pang, R. H. L., Yip, Y. K., Henrikson, D. & Vilček, J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3469–3472.
11. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
12. Blattner, F. R., Blechl, A., Denniston-Thompson, K., Faber, H. E., Richards, J. E., Slightom, J. L., Tucker, P. W. & Smithies, O. (1978) *Science* **202**, 1279–1284.
13. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
14. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 560–580.
15. Ohno, S. & Taniguchi, T. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5305–5309.
16. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
17. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. (1980) *Nature (London)* **283**, 220–224.
18. Huang, A. L., Ostowsky, M. C., Berard, D. & Hager, G. L. (1981) *Cell* **27**, 245–255.
19. Mayo, K. E., Warren, R. & Palmiter, R. D. (1982) *Cell* **29**, 99–108.
20. Brinster, R. L., Chen, H. Y., Warren, R., Sarthy, A. & Palmiter, R. D. (1982) *Nature (London)* **296**, 39–42.
21. Robbins, D. M., Poek, I., Seeburg, P. H. & Axel, R. (1982) *Cell* **29**, 623–631.
22. Weidle, U. & Weissmann, C. (1983) *Nature (London)* **303**, 442–446.
23. Ohno, S. & Taniguchi, T. (1983) *Nucleic Acids Res.* **11**, 5403–5412.
24. Stewart, W. E., II (1979) *The Interferon System* (Springer, New York).
25. Yip, K. K., Pang, R. H. L., Urban, C. & Vilček, J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1601–1605.
26. Farrar, W. L., Johnson, H. M. & Farrar, J. J. (1981) *J. Immunol.* **126**, 1120–1125.
27. Yamamoto, J. K., Farrar, W. L. & Johnson, H. M. (1982) *Cell. Immunol.* **66**, 333–340.
28. Pearlstein, K., Palladino, M. A., Welte, K. & Vilček, J. (1983) *Cell. Immunol.* **80**, 1–9.
29. Gray, P. W. & Goeddel, D. V. (1982) *Nature (London)* **298**, 859–863.
30. Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G. & Fiers, W. (1982) *EMBO J.* **1**, 953–958.
31. Gillis, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. (1983) *Cell* **33**, 717–728.
32. Benerji, J., Olson, L. & Schaffner, W. (1983) *Cell* **33**, 729–740.
33. Moreau, P., Hen, R., Waslyk, B., Everett, R., Gaub, M. P. & Chambon, P. (1981) *Nucleic Acids Res.* **9**, 6047–6068.
34. Weiner, H., König, M. & Gruss, P. (1983) *Science* **219**, 626–631.
35. Banerji, J., Rusconi, S. & Schaffner, W. (1981) *Cell* **27**, 299–308.
36. Khoury, G. & Gruss, P. (1983) *Cell* **33**, 313–314.