Molecular Cloning and Structure of the Human Interleukin-5 Gene*

(Received for publication, April 4, 1987)

Toshizumi Tanabe, Mikio Konishi, Tatsumi Mizuta, Takaumi Noma, and Tasuku Honjo

From the Department of Medical Chemistry, Kyoto University Faculty of Medicine, Kyoto 606, Japan

We isolated the chromosomal gene for human interleukin-5 (IL-5) from human genomic libraries using a cloned human IL-5 cDNA as probe. Nucleotide sequence determination of the IL-5 gene and its flanking regions showed that the gene consisted of four exons and three introns. TATA-like and CAAT-like sequences reside 26 and 73 base pairs, respectively, upstream of the transcription initiation site identified by S1 mapping analysis. The 5'-flanking region of the IL-5 gene has sequences homologous with those of the corresponding regions of the genes for human granulocyte/macrophage colony-stimulating factor, murine IL-4, human interferon-γ, and human IL-2. The intron-exon organization and location of the translation start residue suggest that the IL-5 gene is phylogenetically related to the genes for the granulocyte/macrophage colony-stimulating factor, IL-4, and IL-2, although their amino acid sequences are not significantly conserved.

Several soluble factors are involved in the proliferation and maturation of B lymphocytes (1, 2). cDNAs for two B-cell factors named interleukins-4 and -5 have been cloned and characterized (3-7). The two recombinant factors have several biological activities that seemed to explain most, if not all, of the B-cell factor activities (3-7). IL-4 functions as an immunoglobulin G1 induction factor (3), B-cell-stimulating factor (3, 4), T-cell growth factor (4, 7, 8) and mast cell growth factor (4, 9). IL-5 acts as a T-cell-replacing factor (5), B-cell growth factor II (5), and eosinophil differentiation factor (10). Thus, both IL-4 and IL-5 have multifunctions and a number of target cells. It has been assumed that differentiation and growth factors are distinct molecular entities (11), but IL-4 and IL-5 both stimulate growth and affect maturation under different conditions. The availability of recombinant IL-4 and IL-5 prompted us to re-examine the steps of maturation of B lymphocytes.

Although both IL-4 and IL-5 are classified as lymphokines, which are often synthesized simultaneously by T-cells upon antigen or mitogen stimulation, comparison of the primary structures of six T-cell-derived lymphokines, IL-2, IL-3, IL-4, IL-5, interferon-γ (IFN-γ), and granulocyte/macrophage colony-stimulating factor (GM-CSF), has not shown any strong homology (3-7). In addition, their functions and target cells are various. To trace the phylogenetical relationship among these divergent yet related lymphokine genes, their intron-exon organization must be compared. Isolation of the chromosomal genes for lymphokines is also important for the elucidation of the regulatory mechanism of their expression.

For these purposes, we cloned and characterized the chromosomal gene for human IL-5 and here report its complete nucleotide sequence and organization. Comparison of the chromosomal genes of IL-5 and other lymphokines showed that the IL-5 gene is indeed related phylogenetically to the genes for GM-CSF, IL-4, IL-2, and IFN-γ.

**EXPERIMENTAL PROCEDURES**

**Materials**—Restriction endonucleases, T4 ligase, T4 DNA polynucleotide kinase, T4 DNA polymerase, T4 polynucleotide kinase, S1 nuclease, and the M13 sequence

![Fig. 1. Restriction site map, organization, and sequencing strategy of the human IL-5 gene. a, human IL-5 gene clones and their restriction map. λ12, λ22, and λ38 were isolated from libraries of human fetal liver DNA, human placental DNA, and DNA from human myeloma cell line 266B1, respectively. B, EcoRI; H, HindIII; B, BamHI; b, organization of the IL-5 gene and sequencing strategy. Boxes and lines represent four exons and three introns, respectively. The closed and open boxes indicate coding and untranslated regions, respectively. Horizontal arrows indicate ranges and directions of the sequences read. c, human IL-5 cDNA clone and restriction sites used for isolation of probes. Closed and open boxes indicate coding and untranslated regions, respectively.**

*This work was supported by grants from the Ministry of Education, Science, and Culture of Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J09478.

The abbreviations used are: IL, interleukin; IFN-γ, interferon-γ; GM-CSF, granulocyte/macrophage colony-stimulating factor; bp, base pair; kb, kilobase pair; ATL, human adult T cell leukemia; PIPES, 1,4-piperazinediethanesulfonic acid.
cloned into pUC18 vector. A series of unidirectional deletion mutant HindIII. The two 1.6-kb fragments thus produced were then sub-

method (17).

clones was prepared by digestion of each plasmid with exonucleases by Southern blot hybridization (15) was isolated and digested with of clone A12 which was hybridized with the human IL-5 cDNA probe of clone A12 which was hybridized with the human IL-5 cDNA probe. The probe (10^6 cpm) was denatured

&

Schull. Exonucleases probe. The probe (10^6 cpm, 5 x 10^6 cpm/µg of DNA) was denatured

&

2,745

910

TAC CGA CAG TTT CCT GGT GAC TTA ATT GAA GTT CCG CGC CTC AAA TCC CCG CAG GAA CAG GTA CCG CGA CAA ACG AAA AAG AAC

His His Leu Val Arg Leu

Fig. 2. Nucleotide sequences of the human IL-5 gene and its flanking regions. Translated amino acid residues of the coding sequence are shown below the nucleotide sequence. The asterisk (position 509) indicates the transcription initiation site identified by S1 mapping analysis. Underlined sequences are CAAT-like, TATA-like, and poly(A) signal sequences from 5' to 3', in that order. Inverted repeats with a stem-loop structure (positions 350-383) are shown by two facing arrows. Arrowheads indicate the position of poly(A) attachment.

SI Mapping—Digestion of the 3.2-kg BamHI fragment of the IL-5 gene with Tsai produced a 642-bp fragment containing the 5'-flanking sequence and a part of exon 1, which was then labeled at the 5' end and with [γ-32P]ATP and T4 polynucleotide kinase and used as a probe. The probe (10^6 cpm, 5 x 10^6 cpm/µg of DNA) was denatured and mixed with 19 µg of poly(A) RNA from ATL-2 cells that produce IL-5 (6). Hybridization was done at 39°C for 15 h in a solution containing 80% formamide, 0.4 M NaCl, 1 mM EDTA, 0.04 M PIPES buffer (pH 6.4), and the hybrid obtained was digested with 43 units of S1 nuclease by the method of Berk and Sharp (18). The products were electrophoresed through a 6% polyacrylamide gel containing 7 M urea along with dye DNA sequencing samples of the 3'-half of the HindIII-BamHI fragment of the IL-5 gene. The gel was dried under reduced pressure and autoradiographed.

RESULTS

Cloning of the Human IL-5 Gene—Phage libraries containing human fetal liver DNA, human placental DNA, and DNA of the human myeloma cell line 266B1 were screened with the Psul-PstI fragment of the human IL-5 cDNA containing the 5'-untranslated region and the entire coding region as probe (Fig. 1). Positive clones λ12, λ22, and λ38 were obtained by screening 5 x 10^9 plaques each of DNA libraries of fetal liver, placenta, and myeloma, respectively. Results of restriction

kit were purchased from Takara Shuzo Co., Ltd. (Kyoto, Japan) and Toyobo Co., Ltd. (Osaka, Japan). [α-32P]dCTP (3000 Ci/mmole) and [γ-32P]ATP (5000 Ci/mmole) were obtained from Amersham Corp. and Du Pont-New England Nuclear. Nitrocellulose filters (pore size, 0.45 µm) were purchased from Schleicher & Schuell. Exonucleases HI and VII were purchased from Bethesda Research Laboratories.

Phage Libraries—The Charon 4A phage library containing AluI-PstI-PstI-PstI partial digest of human myeloma DNA contains EcoRI partial digests of human placental DNA and one containing EcoRI partial digest of DNA from human myeloma cell line 266B1 that produces IgE were prepared by Takahashi and Honjo and Nishida et al. (13), respectively. Phages were screened by the method of Benton and Davis (14) with the 515-base pair (bp) PstI-PstI fragment of human IL-5 cDNA (6) as probe.

Nucleotide Sequence Determination—The 3.2-kg BamHI fragment of clone X12 which was hybridised with the human IL-5 cDNA probe by Southern blot hybridization (15) was isolated and digested with HindIII. The two 1.6-kb fragments thus produced were then subcloned into pUC18 vector. A series of unidirectional deletion mutant clones was prepared by digestion of each plasmid with exonucleases HI and VII from either end (16). Nucleotide sequences of inserts of mutant plasmids were determined by the dideoxy chain termination method (17).
enzyme cleavage mapping and Southern blot hybridization with the PstI-PstI fragment of the human IL-5 cDNA showed that the inserts of the three clones isolated overlapped with each other and contained the same 3.2-kb BamHI fragment that hybridized with the IL-5 cDNA (Fig. 1). A smaller fragment of human IL-5 cDNA (the PstI-BamHI fragment) containing the 3’-untranslated region did not hybridize with bands other than the 3.2-kb BamHI fragment, so this fragment probably includes all the exons. We therefore determined its nucleotide sequence by the strategy shown in Fig. 1.

**Nucleotide Sequence of the Human IL-5 Gene**—The nucleotide sequence of the human IL-5 gene is shown in Fig. 2. The gene was found to consist of four exons and three introns by comparison with the cDNA sequence (6). Each intron interrupts the reading frame exactly between codons and contains GT and AG consensus dinucleotides (19, 20) at the 5’ and 3’ boundaries, respectively. Exon 1 (positions 509–696) encoded 48 amino acid residues and contained the 5’-untranslated region. Exon 2 (positions 905–937) and exon 3 (positions 1883–2011) encoded 11 and 43 residues, respectively. Exon 4 (positions 2118–2583) encoded 32 residues and contained the 3’-untranslated region. The lengths of introns 1–3 were 208, 945, and 106 bp, respectively.

**Transcription Initiation Site and Nucleotide Sequence of the 5’-Flanking Region of the Human IL-5 Gene**—To identify the transcription initiation site of the IL-5 gene, S1 mapping was done with a TaqI-TaqI fragment (positions 30–672) labeled at the 5’ end and poly(A+) RNA from ATL-2 cells, from which we have cloned cDNA for IL-5 (6). The size of the protected fragment was 166 bases (Fig. 3), so the 5’ terminus of IL-5 mRNA was at position 509, which corresponds to the 5’ end of the cloned human IL-5 cDNA (6). The primary transcript of the IL-5 gene must be 2075 bases from the transcription initiation site to the beginning of the poly(A) tail of the IL-5 cDNA (position 2583).

The nucleotide sequence of the 5’-flanking region upstream of the transcription initiation site was also determined and characterized. A TATA-like sequence (AATTTG) and a CAAT-like sequence (GGGATTCT) were found 26 and 73 bp upstream of the transcription initiation site, respectively. Another structure of the 5’-flanking region was a stem-loop structure at positions 350–383, in which 14 nucleotides of blocks of inverted repeats, i.e. (5′GATTAAAGATATAAA3′) and (5′TTATTTTTTTAAGA3′), were separated by six nucleotides. Other sequence elements of interest will be described under “Discussion.”

**DISCUSSION**

We determined the total nucleotide sequence of the human IL-5 gene and analyzed its organization. The perfect agreement between the nucleotide sequences of the chromosomal gene exons and the cDNA sequence found previously (6) indicates that no structural change has occurred in the exon sequences of the IL-5 gene of the ATL-2 cell line from which we cloned cDNA for IL-5 (6). ATL-2 cells contain the human T lymphotropic virus I genome and express large numbers of IL-2 receptors (21). IL-2 receptors (21). IL-5, IL-4, IL-2, and GM-CSF genes each contained 1 cysteine residue close to the 5’ ends of the peptide, but only exon 1 had 3 cysteine residues. The IL-5 gene isolated was shown to encode biologically active IL-5 by transfection of cultured cells.

*3 T. Noma and T. Honjo, unpublished data.

We have compared the intron-exon organization of the human genes for IL-5, IL-2 (22), IL-3 (23), GM-CSF (24, 25), IFN-γ (26), IL-1α (27), granulocyte colony-stimulating factor (28), and the murine gene for IL-4 (29), for which the human sequence was not available. The IL-5, IL-4, IL-2, and GM-CSF genes had three introns at roughly similar positions of the polypeptides (Fig. 4). The IL-5, IL-4, IL-2, and GM-CSF genes each contained 1 cysteine residue close to the 5’ end of exons 3 and 4. The IFN-γ gene also had three introns at roughly similar positions of the peptide, but only exon 1 had 3 cysteine residues. The IL-3 gene had four introns, but the relative locations of the first, second, and fourth introns resembled those of the three introns of the IL-5, IL-4, GM-CSF, and IL-2 genes, which may have lost one intron during evolution. The cysteine residue was not found near the 5’ end.
The IFN-γ further upstream. A sequence homologous to one of the T- and IL-2 genes have homology with the IL-5 gene in a region of gaps.

Two sequences are aligned to maximize homology by the introduction of gaps. Numbers above and below the sequences refer to nucleotide positions from the transcription initiation site.

of exon 3 of the IL-3 gene. These comparisons suggest that the IL-5 gene is related phylogenetically to the genes for GM-CSF, IL-4, IL-2, and IFN-γ. However, the IL-5 gene is much less closely related to the IL-3 gene. In contrast, the numbers and relative locations of introns of the IL-5 gene were quite different from those of the IL-1α and granulocyte colony-stimulating factor genes, which are not usually expressed by T-cells.

The 5′-flanking regions of the human IL-5 gene and other lymphokine genes expressed in T-cells (GM-CSF, IFN-γ, IL-2, IL-3 [30], and IL-4) were compared to see whether they shared common regulatory sequences. Short homologous sequences were found here and there; only homologous pairs at similar locations relative to the transcription initiation site are shown in Fig. 5. Homology between the IL-5 and GM-CSF (24, 25) genes was significant immediately upstream of the TATA box. This region also has sequence homology between the human IL-2 and IFN-γ genes (22). The IFN-γ and IL-2 genes have homology with the IL-5 gene in a region further upstream. A sequence homologous to one of the T-cell-specific regulatory sequences (31) was also found in the human IL-5 gene (positions 323–339, TTTAAGGGGGAGAC). The consensus sequence is present at both sides of the enhancer-like regulatory sequence in the IL-2 gene (31), but the IL-5 gene contained only one set of the consensus sequence. It is not known whether these regions are important in the transcriptional regulation of lymphokine genes.

Acknowledgment—We thank K. Hirano for help in preparation of the manuscript.

REFERENCES
1. Dutton, R. W., Falkoff, R., Hirat, J. A., Hoffman, M., Kappler, J. W., Kettman, J. R., Lesley, J. F., and Vann, D. (1971) Prog. Immunol. 1, 305–368
2. Schimpl, A., and Wecker, E. (1972) Nature New Biol. 237, 15–17
3. Noma, Y., Sideras, P., Naito, T., Bergstedt-Lindqvist, S., Azuma, C., Severinson, E., Tanabe, T., Kinashi, T., Matsuda, F., Yaota, Y., and Honjo, T. (1986) Nature 319, 640–646
4. Lee, F., Yokota, T., Osuka, T., Meyerson, P., Villaret, D., Coffman, R., Mosmann, T., Renick, D., Roehm, N., Smith, C., Zlotnik, A., and Arai, K. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2061–2065
5. Kinashi, T., Harada, N., Severinson, E., Tanabe, T., Siders, P., Konishi, M., Azuma, C., Tomininga, A., Bergstedt-Lindqvist, S., Takahashi, M., Matsuda, F., Yaota, Y., Takatou, K., and Honjo, T. (1986) Nature 324, 70–75
6. Azuma, C., Tanabe, T., Konishi, M., Kinashi, T., Noma, T., Matsuda, F., Yaota, Y., Takatou, K., Hammarström, L., Smith, C. I. E., Severinson, E., and Honjo, T. (1986) Nucleic Acids Res. 14, 9149–9158
7. Yokota, T., Osuka, T., Mosmann, T., Banchesoi, J., DeFrance, T., Blanchard, D., De Vries, J. E., Lee, F., and Arai, K. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 5894–5898
8. Severinson, E., Naito, T., Tokumoto, H., Hama, K., and Honjo, T. (1987) Eur. J. Immunol. 17, 67–72
9. Hamaguchi, Y., Kanakura, Y., Fujita, T., Takeda, S., Nakano, T., Tanui, S., Honjo, T., and Kitamura, Y. (1987) J. Exp. Med. 165, 268–273
10. Sanderson, C. J., O'Garra, A., Warren, D. J., and Klaus, G. B. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 437–440
11. Howard, M., Nakanishi, K., and Paul, W. E. (1984) Immunol. Rev. 75, 185–210
12. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., and Efstratiadis, A. (1978) Cell 15, 687–701
13. Nishida, Y., Miki, T., Hisajima, H., and Honjo, T. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 3833–3837
Human IL-5 Gene

14. Benton, W. D., and Davis, R. W. (1977) *Science* **196**, 180–182
15. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517
16. Yanish-Perron, C., Vieira, J., and Messing, J. (1985) *Gene (Amst.)* **33**, 103–119
17. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5469
18. Berk, A. J., and Sharp, P. A. (1977) *Cell* **12**, 721–732
19. Breathnach, R., Benoist, C., O’Hare, K., Gannon, F., and Cham- b, P. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75**, 4853–4857
20. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., and Steitz, J. A. (1980) *Nature* **283**, 229–224
21. Sabe, H., Kondo, S., Shimizu, A., Tagaya, Y., Yodoi, J., Koba- yashi, N., Hatanaka, M., Maeda, M., Noma, T., and Honjo, T. (1984) *Mol. Biol. Med.* **2**, 379–396
22. Fujita, T., Takakura, C., Matsui, H., and Taniguchi, T. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 7437–7441
23. Yang, Y.-C., Ciarletta, A. B., Temple, P. A., Chung, M. P., Kovacs, S., Witek-Giannotti, J. S., Leary, A. C., Kris, R., Donahue, R. E., Wong, G. G., and Clark, S. C. (1986) *Cell* **47**, 3–10
24. Kaushansky, K., O’Hare, P. J., Berkner, K., Segal, G. M., Hagen, P. S., and Adamson, J. W. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3101–3105
25. Miyatake, S., Otsuka, T., Yokota, T., Lee, F., and Arai, K. (1985) *EMBO J.* **4**, 2561–2568
26. Gray, P. W., and Goeddel, D. V. (1982) *Nature* **298**, 859–863
27. Furutani, Y., Notake, M., Fukui, T., Ohue, M., Nomura, H., Yamada, M., and Nakamura, S. (1986) *Nucleic Acids Res.* **14**, 3167–3179
28. Nagata, S., Tsuichiya, M., Asano, S., Yamamoto, O., Hirata, Y., Kubota, N., Oheda, M., Nomura, H., and Yamazaki, T. (1986) *EMBO J.* **5**, 575–581
29. Otsuka, T., Villaret, D., Yokota, T., Takebe, Y., Lee, F., Arai, N., and Arai, K. (1987) *Nucleic Acids Res.* **15**, 333–344
30. Miyatake, S., Yokota, T., Lee, F., and Arai, K. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 316–320
31. Fujita, T., Shibuya, H., Ohashi, T., Yamanishi, K., and Taniguchi, T. (1986) *Cell* **46**, 401–407