Structure of the Human Interleukin-2 Receptor Gene

Warren J. Leonard, Joel M. Depper, Minoru Kanehisa
Martin Krönke, Nancy J. Peffer, Penny B. Svetlik
Margery Sullivan, Warner C. Greene

Abstract. *The gene encoding the human interleukin-2 (IL-2) receptor consists of 8 exons spanning more than 25 kilobases on chromosome 10. Exons 2 and 4 were derived from a gene duplication event and unexpectedly also are homologous to the recognition domain of human complement factor B. Alternative messenger RNA (mRNA) splicing may delete exon 4 sequences, resulting in a mRNA that does not encode a functional IL-2 receptor. Leukemic T cells infected with HTLV-I and normal activated T cells express IL-2 receptors with identical deduced protein sequences. Receptor gene transcription is initiated at two principal sites in normal activated T cells. Adult T cell leukemia cells infected with HTLV-I show activity at both of these sites, but also at a third transcription initiation site.*

Antigen- or mitogen-induced activation of resting T cells stimulates the synthesis and secretion of interleukin-2 (IL-2 or T cell growth factor) (*1*). These activation signals also induce the expression of specific membrane receptors for IL-2 on another, possibly overlapping, subpopulation of T cells (*2*). In the presence of IL-2, the cells with IL-2 receptors proliferate, resulting in the expansion of T cell populations subserving helper, suppressor, or cytotoxic effector functions (*3*). Both IL-2 and IL-2 receptors therefore have a critical role in the evolution of specialized T cell populations necessary to mount a normal T-cell immune response. Furthermore, recent studies suggest that IL-2 receptors may also play a role in normal B cell growth and differentiation (*4*, *5*).

IL-2 receptors have been characterized as 55,000-dalton glycoproteins that contain *N*- and *O*-linked carbohydrate (*6*, *7*) and are both sulfated (*8*) and phosphorylated (*8*, *9*). While most T cell leukemias do not display IL-2 receptors, adult T cell leukemia (ATL) cell lines, all of which are infected with HTLV-I, uniformly express large numbers of IL-2 receptors (*10*). Select ATL cell lines, such as HUT-102B2, express receptors abnormal in size as evaluated by migration on sodium dodecyl sulfate (SDS) polyacrylamide gels (*7*, *8*, *10*). It is possible, although it has not yet been proved, that these receptors may be involved in the uncontrolled growth of these leukemic cells.

Complementary DNA's (cDNA's) for the human IL-2 receptor have been isolated from libraries prepared with messenger RNA (mRNA) from HTLV-I infected T cell lines (*11–13*). The DNA sequence predicts that the mature protein consists of 251 amino acids. This protein, on the basis of hydrophobicity plots (Fig. 2) (*14*), contains a 219-residue extracellular domain, a 19-residue hydrophobic transmembrane domain, and a cytoplasmic domain only 13 amino acids long. The human IL-2 receptor has been identified as being encoded by a single gene located on chromosome 10, bands p14→15 (*15*). We now present the sequence and structure of this growth factor receptor gene, and provide new insights into the alternative mRNA splicing and the multiple mRNA species that originate from this single gene in both normal and HTLV-I infected leukemic T cells.

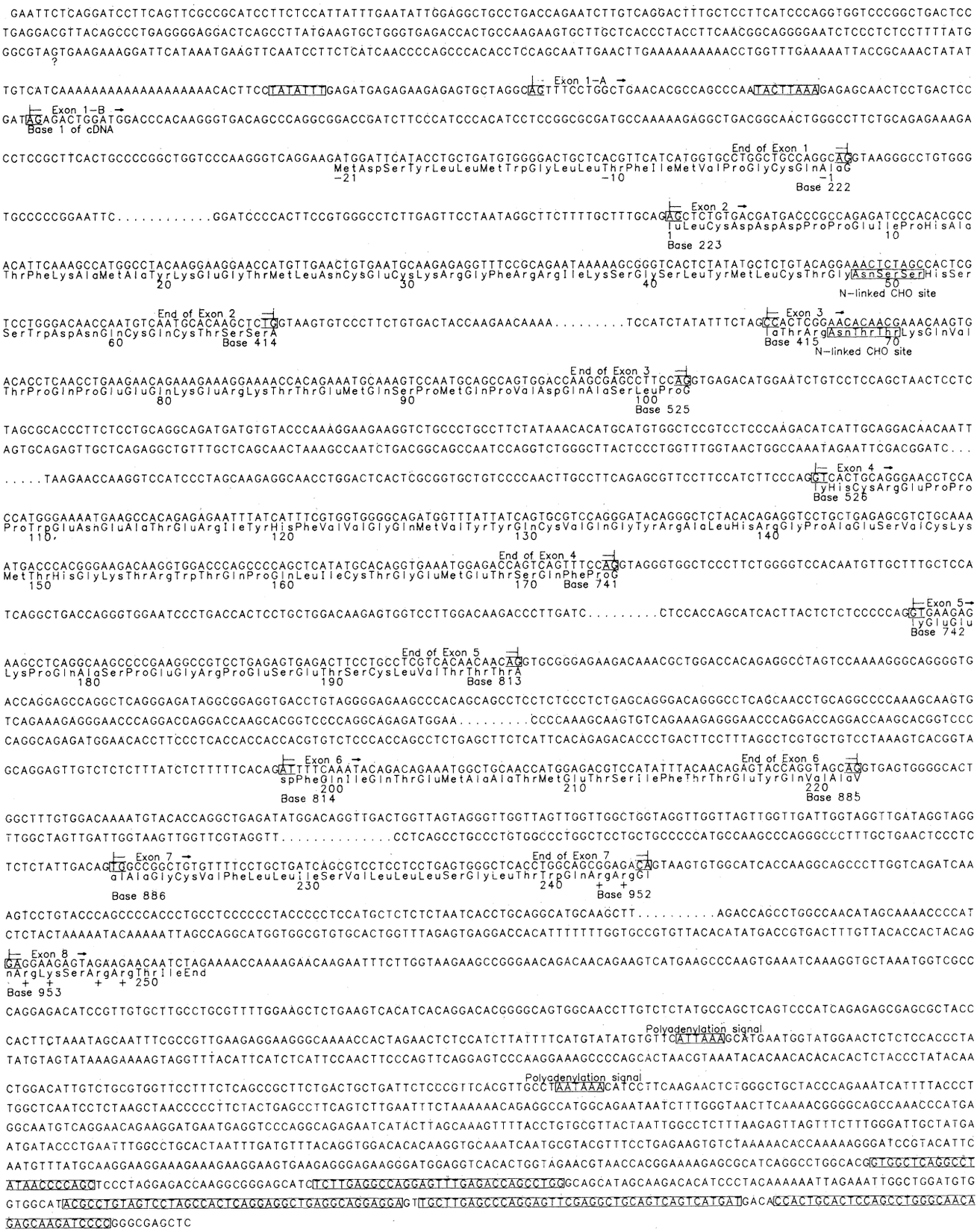
The IL-2 receptor gene: Sequence of exons and exon-intron splice junctions. ^{32}P-labeled IL-2 receptor cDNA hybridizes to five Eco RI fragments of approximately 10 kb, 6 kb, 2.4 kb, 1.4 kb, and 0.7 kb on genomic Southern blots (*11*, *16*). In order to further elucidate the structure of the IL-2 receptor gene, we isolated seven IL-2 receptor genomic phage clones (λ-IL2R-A, -B, -C, -D, -E, -F, and -G) from two different libraries (*17*). Together these clones contained all five of these Eco RI fragments, which were subcloned into pBR322 or pUC13.

The DNA sequence of each of the exons and exon-intron splice junctions of the IL-2 receptor gene was determined by the dideoxy chain termination method. We compared these sequences to the sequence of the IL-2 receptor cDNA clone, pIL2R3 (*11*), using the program of Queen and Korn (*18*) on an IBM system 370 computer. These sequence data revealed that the IL-2 receptor gene consists of eight exons and seven introns. The sequences of the exons and the corresponding exon-intron splice junctions are presented in Fig. 1.

Analysis of the exons revealed that the physical structure of the gene generally correlated with the peptide domains of the receptor, such as the transmembrane and cytoplasmic domains, as predicted by a Kyte-Doolittle hydrophobicity plot (Fig. 2). The 5′ untranslated and signal peptide sequences were contained within the first exon. The 5′ end of exon 2, which is located at least 15 kb from exon 1, corresponds to the first amino acid of the mature receptor protein. Exons 2 and 3 each encode one of the two *N*-linked glycosylation sites and are rich in serine or threonine, suggesting that they may also contain *O*-linked carbohydrate addition sites. Furthermore, exon 2 shares approximately 25 percent amino acid homology, including conservation of the cysteine, proline, and tryptophan residues, with exon 4, suggesting a gene duplication event within this region. This homology is indicated by the dot matrix analysis presented in Fig. 3. Similar internal homology was also detected within a mouse IL-2 receptor cDNA, suggesting that this internal gene duplication occurred at least 50 million years ago prior to the murine radiation (*19*). Exons 5 and 6 correspond to the region of the receptor located immediately exterior to the cell membrane and, like exon 2, exons 5 and 6 encode clusters of serines or threonines and thus contain potential *O*-linked carbohydrate addition sites. In this sense, the IL-2 receptor is somewhat

W. J. Leonard is in the Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, Bethesda, Maryland 20892. J. M. Depper is in the Department of Medicine, University of Michigan School of Medicine, Ann Arbor 48109. M. Kanehisa is at the Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan. M. Krönke is at the Klinische Arbeitsgruppe, Gosslerstr. 10 d, 3400 Göttingen, West Germany. M. Sullivan is at Genex Corporation, Rockville, Maryland 20850. N. J. Peffer, P. B. Svetlik, and W. C. Greene are in the Metabolism Branch, National Cancer Institute, Bethesda 20892. Please address correspondence to W. J. L., Building 18T, Room 101, National Institutes of Health.

analogous to the low-density lipoprotein (LDL) receptor, which also contains multiple *O*-linked glycosylation sites located immediately external to the cell membrane (*20*, *21*). Exon 7 contains most of the transmembrane domain. Exons 7 and 8 encode the cytoplasmic domain, which is only 13 amino acids long but contains six positively charged amino acids and thus probably serves a cytoplasmic anchoring function. Exon 8 also encodes the 3′ untranslated region including at least three polyadenylation signals. Unexpectedly, the 3′ untranslated region also contains multiple copies of the Alu family of repetitive DNA sequences, located between the second and third polyadenylation sites. These



sequences are present in the longest IL-2 receptor cDNA identified (pIL2R3) (*11*), and therefore occur in mature mRNA. Similar Alu repetitive sequences have been identified in the 3′ untranslated region of the LDL receptor (*20*).

Alternative mRNA splicing of exon 4. IL-2 receptor cDNA's lacking a 216-bp segment within the protein coding region were previously isolated (*11*, *13*). S1 nuclease protection assays have demonstrated the presence of mRNA's corresponding to this abnormal cDNA in both normal activated T cells (*22*) and HTLV-I infected leukemic T cell lines (*16*). When expressed in either COS-1 cells or L cells, this alternatively spliced cDNA does not appear to direct the expression of membrane receptors capable of binding either IL-2 or a monoclonal antibody to the IL-2 receptor (anti-Tac) (*11*, *23*). It is now evident that the missing 216-bp segment corresponds precisely to exon 4, and that alternative posttranscriptional splicing is responsible for producing the mRNA's in which exon 4 is deleted and exons 3 and 5 are joined to each other.

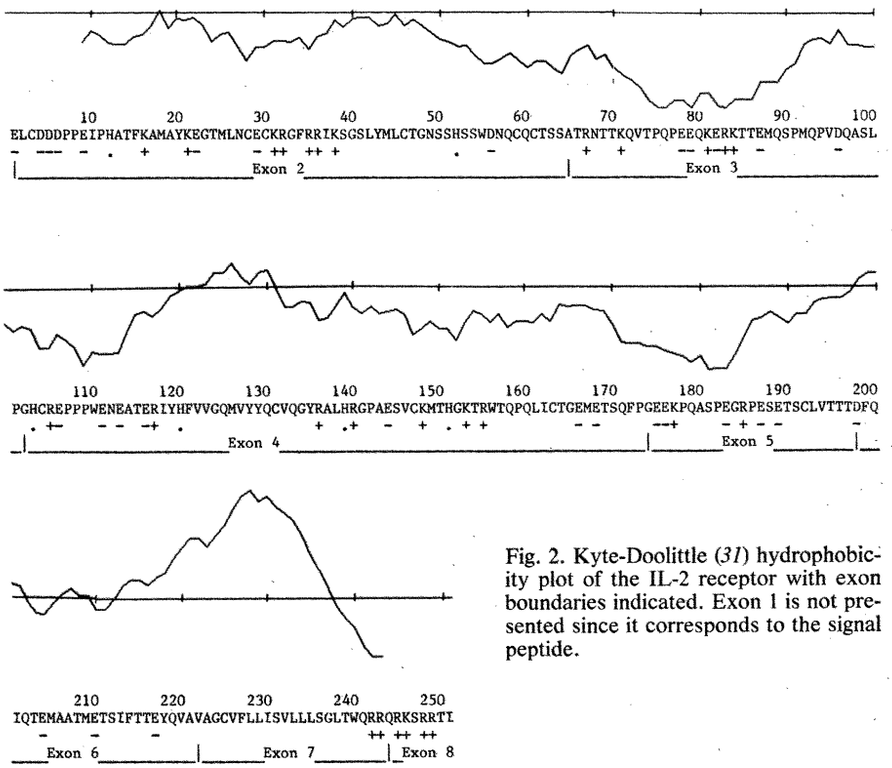


Fig. 2. Kyte-Doolittle (*31*) hydrophobicity plot of the IL-2 receptor with exon boundaries indicated. Exon 1 is not presented since it corresponds to the signal peptide.

Normal and ATL IL-2 receptors share identical amino acid sequence; homology to human complement factor B. Comparison of the exon coding sequences, which were derived from normal donor DNA, to the sequence of the HUT-102B2 derived cDNA revealed no differences; thus, both the normal and HUT-102B2 leukemic IL-2 receptor sequences are identical, even though the mature proteins differ by approximately 5000 daltons, on the basis of their migration on SDS polyacrylamide gels (see below). Comparison of the deduced amino acid sequences to the National Biomedical Research Foundation Database revealed no significant homologies. However, when we used a new program (*24*) which converts all DNA sequences in the GenBank database to protein sequences, we demonstrated that exon 4 shares approximately 30 percent amino acid homology with the Ba fragment of human complement factor B (*25*). This homology involves the recognition (binding) domain, but not the serine protease domain, of this complement protein. The recognition domain of factor B has been described as having DNA binding properties. The homologous regions are indicated in Fig. 4. The LDL receptor shares significant homology with the recognition domains of coagulation factors IX and X, which also participate in a proteolytic cascade (*21*). By comparison of alignment scores, the homology of the LDL receptor to factors IX and X is approximately 9.4 standard deviation units; similarly, the homology of the IL-2 receptor to complement factor B is 9.659 standard deviation units, indicating highly significant homology.

The physical map of the IL-2 receptor gene is shown in Fig. 5. The total gene spans a minimum distance of 25 kb. At present, a more precise size cannot be given because of the uncertainty regarding the size of the first intron. While phage λ-IL2R-C contains exon 1, it does not extend sufficiently in the far 3′ direction to overlap any of the other phage clones. On the basis of restriction mapping of λ-IL2R-C, the first intron is at least 15 kb in length. In studies of the murine IL-2 receptor, Miller and Germain have isolated two mouse IL-2 receptor cosmid clones which lack exon 1 and contain either 15 kb or 30 kb of sequence that is located 5′ to exon 2 (*26*). Thus, if the human gene parallels that of the mouse, it is possible that the first intron is much longer than 15 kb, and the total gene might be considerably larger than the minimal estimate provided. A heteroduplex analysis of phage clone λ-IL2R-F hybridized to the full length IL-2 receptor cDNA is shown in Fig. 6, A and

Fig. 1. Sequences of exons and exon-intron junctions for the human IL-2 receptor. DNA fragments were subcloned into M13 bacteriophage, sequences were obtained by the dideoxy chain termination method (*29*), as previously described (*11*), and compared to the sequence of the HUT-102B2 cDNA pIL2R3 (*11*). (M13 sequencing reagents were from Pharmacia P-L Biochemicals. [α−^{32}P]dATP, 400 Ci/mmol, was obtained from Amersham.) In the sequence marked "promoter region plus exon 1," the "TATAAA" sequences are boxed in the sequence 5′ to exon 1. The G underscored with a "?" symbol on one sequencing gel appeared to be a C. Left brackets indicate the locations of the two known transcription initiation sites and the 5′ ends of each exon, while right brackets identify the 3′ ends of each exon. Under the left and right brackets, base numbers corresponding to the IL-2 receptor cDNA sequence (*11*) are presented. The numbering scheme omits the previously included 5′ Eco RI site, and poly(C) sequences added during cDNA cloning; therefore base 1 corresponds to transcription initiation site "A" (Figs. 4 and 5). In each case the numbers refer to the first and last base of each exon. Amino acids are numbered every ten residues. The conventional donor and acceptor splice sequences were used to determine exon boundaries; thus, some amino acid three-letter codes are divided between two sequential exons. The dotted areas in Fig. 1 indicate regions of introns not sequenced. The lengths of these regions may be deduced from Fig. 5. Exon 1-A and 1-B correspond to transcription initiation sites A and B. The two *N*-linked carbohydrate addition sites are identified. The sequence of exon 8 as presented includes the first two of three polyadenylation signals. The location of these sequences are identified in exon 8. The repetitive Alu sequences in exon 8 are identified in a series of boxes at the end of the sequence presented in the figure. Positively charged residues in the cytoplasmic domains of exons 7 and 8 are identified with "+" symbols. The sequences were prepared for publication with a DEC10 computer and CALCOMP plotter with the assistance of Marvin Shapiro using the DNA Draw program (*30*).

B. As clone F contains exons 2 through 7, this electron micrograph pictorially depicts the gene except for the first exon and part of the first intron. The IL-2 receptor gene also contains many repetitive intronic DNA sequences, as indicated by the multiple short regions of hybridization within introns (Fig. 5, C and D).

Multiple transcription initiation sites. We next analyzed potential important regulatory regions present 5′ to exon 1. We first performed primer extension experiments to determine whether or not the 5′ end of the cDNA corresponded to the 5′ end of all of the IL-2 receptor mRNA's. A synthetic oligonucleotide probe with the sequence 5′-CCGCCTG-GGCTGTCACCCTTGTGGG-3′, corresponding to bases 37 to 61 of the sequence of pIL2R3 (*11*), was end-labeled with [^{32}P]ATP (adenosine triphosphate) and polynucleotide kinase, and hybridized to mRNA isolated from HUT-102B2 cells or peripheral blood T cells activated with phytohemagglutinin (PHA) or phorbol 12-myristate 13-acetate (PMA). This labeled annealed primer was then extended with reverse transcriptase. Two intense signals were identified both with mRNA from normal activated T cells and HTLV-I infected HUT-102B2 cells (Fig. 7). The smaller of these transcription initiation sites corresponded to the end of the cDNA's previously sequenced [the published sequence (*11*) of pIL2R3 included an added Eco RI linker and poly(C) tail (polycytidylate) at the 5′ end; base 23, not base 1, of that sequence (*11*), is therefore the true 5′ end of the cDNA]. On a sequencing gel, the second band (B) was determined to be located an additional 58 bases 5′ to band A. In addition, as shown in Fig. 7, A and B, a third fainter signal (labeled C) extending approximately 250 bases from the end of the cDNA was also identified with HUT-102B2 mRNA but not with mRNA from normal activated T cells.

In order to clarify whether each of these bands in fact corresponded to mature mRNA's with different 5′ termini as opposed to one large mRNA with regions of secondary structure not permitting efficient reverse transcription, we performed S1 nuclease protection assays. These assays also provided an independent and sensitive method to evaluate whether band C was present only in HTLV-I infected HUT-102B2 cells and not present in PHA activated T lymphoblasts. In these experiments, the 733-bp fragment containing exon 1 and upstream sequences was subcloned into M13 phage in the proper orientation, and a single-stranded ^{32}P-labeled probe, complementary to mRNA, was synthesized. This probe was hybridized to mRNA from various cell types, and S1 nuclease digestion was performed. Two major protected species which corresponded to the two major bands identified in the primer extension experiments were identified (Fig. 8). In addition, a third, larger but less intense, band (labeled C) was identified with HUT-102B2 mRNA, but not with RNA from normal activated T cells. In view of this putative extra transcription site in HUT-102B2 cells, we next analyzed RNA from four other HTLV-I infected T cell lines and other T and B lymphoid cell lines (Fig. 7B). The intense band at 733 bases represents the full-length probe, which is identified even in the absence of mRNA (lane marked control) since, as part of excision from M13 phage, the input probe was digested at both 5′ and 3′ ends and therefore is contaminated by an equal quantity of nonlabeled DNA which protects the full-length probe. The two major bands at approximately 220 and 280

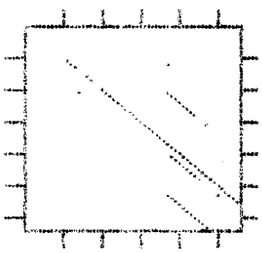


Fig. 3. Dot matrix analysis comparing IL-2 receptor exon 2 on the abscissa (amino acids 1 to 64) with exon 4 on the ordinate (amino acids 102 to 173) using the Mutation Data Matrix (250 PAM's), minimum score = 12, and window size = 20. Alignment scores were calculated and compared to 1000 random sequences. The significance of the homology in standard deviation units was 9.351.

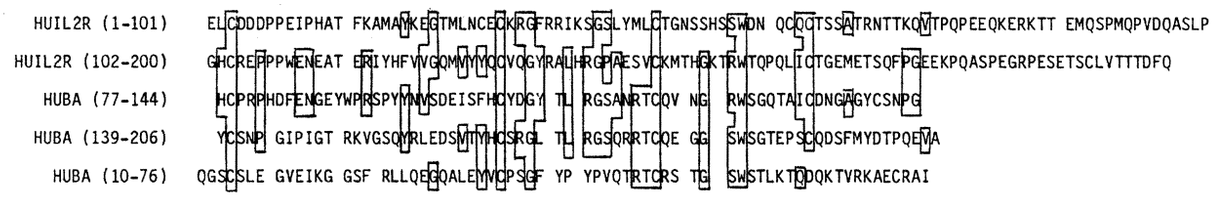


Fig. 4. Amino acid sequence comparison between the IL-2 receptor and the Ba fragment of human complement factor B. Shared amino acids are boxed. Local homologies as high as 30 percent were identified. HUIL2R: sequence of the human IL-2 receptor; HUBA: sequence of human complement factor B, Ba fragment. Alignment scores were calculated for HUIL2R 102 to 173 (exon 4) versus HUBA 77 to 144 and compared to the average distance of 1000 random sequences. The significance of the homology in standard deviation units was 9.659 (*24*).

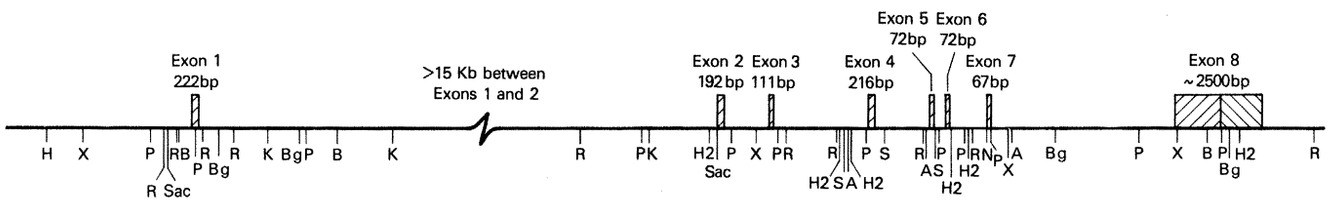


Fig. 5. Map of the IL-2 receptor gene. The gene spans more than 25 kb. DNA from IL-2 receptor genomic phage clones was digested with various restriction enzymes, subjected to electrophoresis on agarose gels, transferred to nitrocellulose paper (either Schleicher & Scheull or Millipore), and hybridized with nick-translated or random-primed ^{32}P-labeled DNA fragments (*32*). These DNA fragments were derived either from genomic clones or from IL-2 receptor cDNA's. Exons are indicated by boxes with crosshatches. The part of exon 8 that was sequenced corresponds to the left half of the box. The right half of the box indicates the estimated distance to the third polyadenylation signal. The size of intron 1 is greater than 15 kb. Enzymes are abbreviated as follows: A, Ava I; B, Bam HI; Bg, Bgl I; H, Hind III; H2, Hinc II; K, Kpn I; N, Nae I; P, Pst I; R, Eco RI; S, Stu I; Sac, Sac I; X, Xba I. The map of most of intron 1 is unknown. Otherwise all Eco RI sites in the gene are indicated. Other restriction enzymes may have more sites than are presented.

bases correspond to bands A and B, as seen in Fig 7A. The band at approximately 410 bases corresponds to the band C of Fig. 7A. This band was identified in HUT-102B2 cells and in four additional ATL cell lines (PLP6, C5MJ, MJ, and C91PL) but not in normal activated T cells or induced JURKAT cells. Of interest, this third band was also detected in the EBV (Epstein-Barr virus)–transformed B cell line, 5B4, after induction of these cells with purified IL-2. As previously reported (*5*), IL-2 induction of these cells results in upregulation of the number of IL-2 receptors expressed. Also seen in Fig. 7, A and B, are two other faint bands at approximately 250 and 310 bases. No obvious TATAAA sequence in the 733-base fragment corresponds to these bands; however, we cannot exclude the possibility that these represent additional transcription initiation sites.

Together, the primer extension and S1 nuclease experiments indicate that there are two principal transcription initiation sites (indicated by the left brackets in Fig. 1 located under the labels exon 1-A and exon 1-B) and thus two promoters governing IL-2 receptor gene expression in normal activated T cells. Both of these promoters are contained within the 733-bp fragment. Each of these promoters is associated with the same single open reading frame, and both promoters appear to be utilized with the same frequency. In addition, in each of five ATL cell lines and one EBV-transformed B cell line expressing IL-2 receptors, there is evidence for a third promoter and corresponding transcription initiation site located 5′ to the 733-base fragment used in the S1 nuclease experiments.

Last, we attempted to demonstrate directly that the region upstream from exon 1 in fact contained functional promoter activity. We prepared a 1.4-kb Pst I fragment which included genomic sequences 5′ to the first exon and the first 105 bp of the cDNA (but stopping 5′ to the ATG translation start site). This fragment was treated with T4 DNA polymerase to generate blunt ends, Hind III linkers were added, and the fragment was cloned in both orientations into the Hind III site of $pA_{10}CAT_2$ (after removal of SV40 promoter and enhancer elements from the vector). These constructs placed the putative IL-2 receptor promoter a short distance upstream from the chloramphenicol acetyltransferase (CAT) gene. These plasmids and pSV2CAT as a positive control were transfected into JURKAT T cells with the use of DEAE dextran, and promoter activity was assessed by measurement of

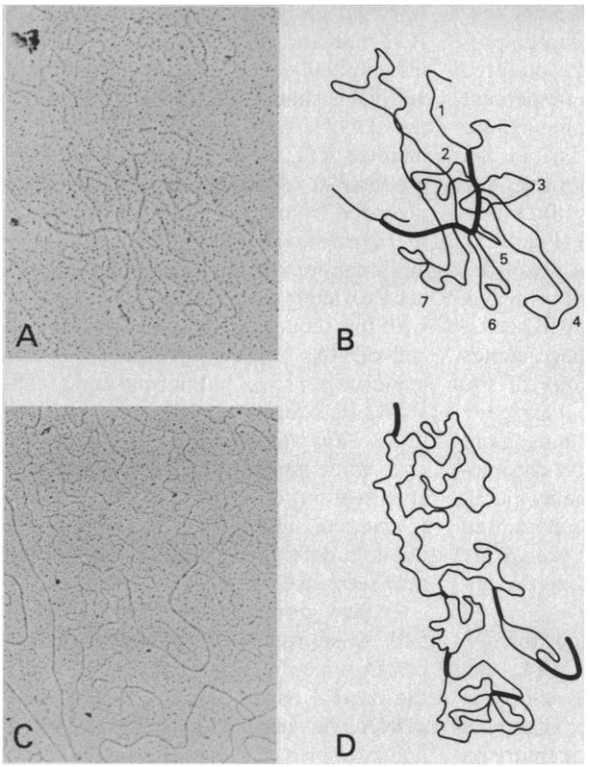


Fig. 6. Heteroduplex (A) and diagram of heteroduplex (B) of λ-IL2R-F to full-length IL-2 receptor cDNA, indicating that the phage clone contains seven exons (all but exon 1). The introns are numbered (B). Some of the numerous repetitive DNA sequences (C) and diagram (D) contained within the gene are indicated by the alternate regions of double-stranded and single-stranded DNA. The DNA from IL-2 receptor genomic phage clones, wild-type λ phage, and IL-2 receptor cDNA were suspended at 2.5 to 5 μg/ml in 0.1*M* tris, 0.01*M* EDTA, *p*H 8.5, 50 percent formamide, heated for 10 minutes at 85°C, and then incubated overnight at room temperature. Photographs of hybridized nucleic acids were measured with a Numonics digitizer.

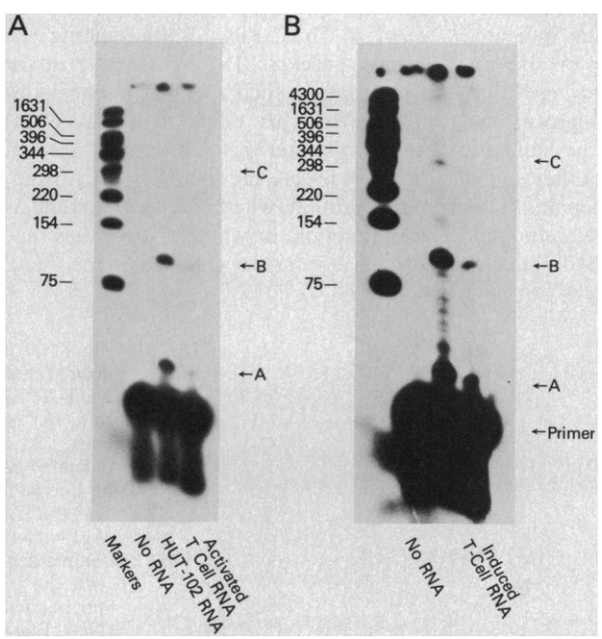


Fig. 7. (A) Primer extension study suggesting three IL-2 receptor gene transcription initiation sites (bands A, B, and C) in HUT-102B2 cells but only two sites (bands A and B) in normal activated T cells. Indicated on the left are molecular weight markers. Primer extension was performed without RNA as a control (second lane) and with HUT-102B2 mRNA (third lane) or activated T cell RNA (fourth lane). (B) Longer exposure of the experiment shown in (A). The primer extension studies were performed approximately as described (*13*). A synthetic oligonucleotide [see (*34*)] corresponding to bases 37 to 61 of the sequence of pIL2R3 was end-labeled with [γ-^{32}P]ATP with the use of polynucleotide kinase (New England Biolabs). After extraction with a mixture of phenol and $CHCl_3$ and precipitation with ethanol, the labeled oligonucleotide was resuspended in water. The labeled oligonucleotide (500,000 count/min) was co-precipitated with 20 μg of total cellular RNA from either HUT-102B2 cells or normal T cells activated with PHA and PMA. This sedimented material was dissolved in 5.5 μl of 10 m*M* tris, *p*H 8.3, 1 m*M* EDTA, and 4.5 μl of 2*M* KCl, heated at 65°C for 5 minutes, and then hybridized at 50°C for 1 hour. The reaction mixture was adjusted to 60 μl containing 50 m*M* tris, *p*H 8.3, 140 m*M* KCl, 10 m*M* $MgCl_2$, 5 m*M* DTT, actinomycin D (100 μg/ml), bovine serum albumin (100 μg/ml), 0.16 m*M* each of the nucleotides dATP, dCTP, dGTP, and dTTP, and 10 units of AMV reverse transcriptase, and incubated at 42°C for 1 hour. The reaction was stopped by adding the sodium salt of EDTA (to 20 m*M*); the mixture was then extracted with phenol-$CHCl_3$ and precipitated with ethanol. The final pellet was dissolved in 100 m*M* tris-borate, *p*H 8.3, 1 m*M* EDTA, and 80 percent formamide, and then subjected to electrophoresis on a 6 percent polyacrylamide gel, 8*M* urea, 1× TBE gel (133 m*M* tris, 80 m*M* boric acid, 2.8 m*M* sodium EDTA).

acetylation of the chloramphenicol substrate. The putative IL-2 receptor promoter region when correctly oriented contained promoter activity (4 percent conversion of [^{14}C]chloramphenicol to acetylated forms). In contrast, in the reverse orientation, this fragment had little or no promoter activity (0.28 percent conversion of substrate) (Fig. 9).

Using IL-2 receptor cDNA's isolated from a library prepared with mRNA from HTLV-I infected HUT-102B2 cells, we have isolated genomic phage clones encompassing all of the exons of the normal IL-2 receptor gene. This gene consists of eight exons spanning more than 25 kb of DNA on human chromosome 10, bands p14 → 15. The availability of these phage clones has permitted us to address a number of previously unanswered questions regarding the human IL-2 receptor.

1) As assessed by migration on SDS polyacrylamide gels, some ATL cell lines, including HUT-102B2 cells, express IL-2 receptors that differ slightly in apparent size from the IL-2 receptors on normal activated T cells. However, other studies have demonstrated no apparent differences in the size of primary translation product or precursor forms of these receptors in normal and ATL cells. Biosynthetic labeling studies indicated that the difference in receptor size appeared to be generated during Golgi-associated posttranslational processing. Thus, it was possible that the leukemic protein coding region might differ slightly from normal, perhaps by a point mutation which would change the amino acid sequence and alter carbohydrate, sulfate, or phosphate addition. Although IL-2 receptor cDNA's had been isolated independently by three groups (*11–13*) and each share identical sequences, all the cDNA's were derived from HTLV-I infected ATL T cell lines, and two of the groups (*11*, *13*) had used cDNA libraries prepared from HUT-102B2 cells. It was therefore essential to sequence the exons and exon-intron junctions of the normal gene to determine identity or nonidentity with the sequence from the HUT-102B2 cell cDNA. The data presented in this manuscript prove that normal and HUT-102B2 IL-2 receptor sequences are identical. Thus, the difference in receptor size, while generated posttranslationally and involving diminished addition of sialic acid and sulfate (*8*), is not apparently due to a difference in the primary amino acid sequence.

2) Two groups had previously identified a shorter IL-2 receptor related cDNA which was missing 216 bp within the protein coding region (*11*, *13*). This cDNA was initially hypothesized (*11*), and later proved by S1 nuclease studies, to be the product of alternative mRNA splicing in both ATL cells (*16*) and normal activated T cells (*22*). Genomic sequence data now indicates that this 216-bp segment corresponds precisely to exon 4. Thus, the mRNA splicing enzymes process IL-2 receptor primary transcripts in such a way as to occasionally delete this exon. It remains unknown whether this alternatively spliced mRNA directs the synthesis of a functional protein, whether it serves a regulatory function, or whether it perhaps represents a recurrent cellular "error," and serves no physiological role. Transfection of this alternatively spliced cDNA in either COS-1 cells or L cells does not result in display of receptors capable of binding either IL-2 or anti-Tac.

3) Since IL-2 receptors in ATL cells are expressed constitutively at a high level (*10*), we considered possible differences in gene structure which could be involved in this deregulated receptor expression. Genomic Southern blots with DNA derived from normal and ATL cells have provided no evidence for receptor gene rearrangement or amplification (*16*). Further, karyotypic studies have not revealed consistent translocations involving the IL-2 receptor locus on chromosome 10 in this leukemia. Thus, neither derangements in IL-2 receptor gene structure nor viral promoter insertion upstream to the IL-2 receptor gene appear to account for the high level receptor expression encountered in this leukemia.

Since it was possible that virally encoded or induced proteins modulated IL-2 receptor expression, we next investigated potential differences in IL-2 receptor promoter structure or function in ATL and normal T cells. We assumed that the sequence 5′ to exon 1 would contain the IL-2 receptor gene promoter. However, initial study of the sequence did not reveal a canonical TATAAA sequence. It was therefore possible that the promoter was contained on an upstream exon that had not yet been identified. We addressed this issue using primer extension and S1 nuclease studies. Primer extension studies revealed two major signals, one corresponding to the

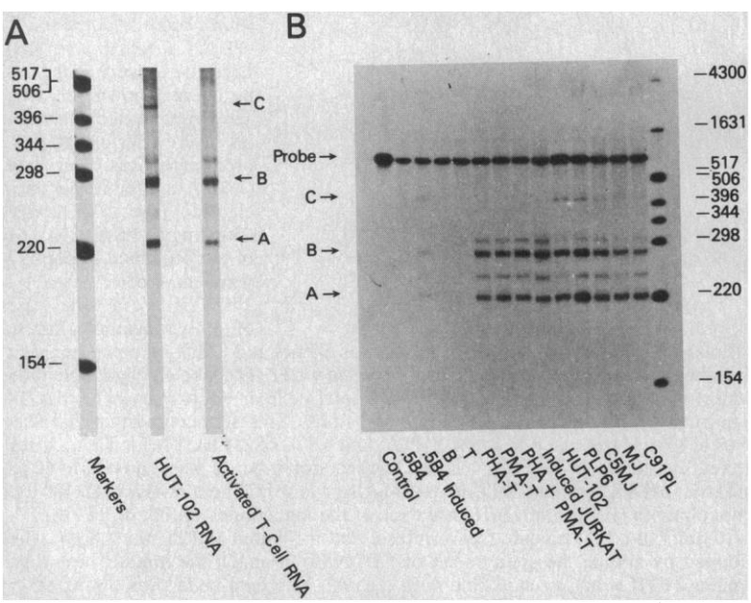


Fig. 8. (A) S1 nuclease studies suggesting three IL-2 receptor gene transcription initiation sites (bands A, B, and C) in HUT-102B2 cells, which constitutively express IL-2 receptors, but only two (bands A and B) in normal activated T cells. (B) S1 nuclease experiment with RNA's from the indicated cell lines. To perform the S1 experiments, the 733-bp Eco RI fragment containing the promoter region and exon 1 of the IL-2 receptor gene (see Fig. 1) was subcloned into M13 bacteriophage. Uniformly labeled single-stranded DNA complementary to IL-2 receptor mRNA was synthesized with the use of the 15-base M13 primer (Pharmacia P-L Biochemicals) and the large fragment of DNA polymerase I (New England Biolabs) in the presence of [α-^{32}P]-labeled dATP or dCTP (400 or 3000 Ci/mmol; Amersham). The DNA was then digested with enzymes (Eco RI, Bam HI, and Hind III), cutting 3′ or both 3′ and 5′ to the insert, and subjected to electrophoresis on a 6 percent acrylamide, 8*M* urea gel in 1× TBE buffer. The labeled fragment (and the complementary unlabeled strand) was recovered by electroelution, extracted with phenol-chloroform, and precipitated with ethanol in the presence of 50 μg of transfer RNA. The purified DNA was denatured in 70 percent formamide, and hybridized for 16 hours at 50°C in a reaction volume of 50 μl (10^6 cpm/ml) to mRNA from various lymphoid cells, including resting and activated T cells, B cells, and HTLV-I infected ATL cell lines. The DNA-RNA hybrids were then digested with S1 nuclease (5000 units; Boehringer Mannheim) for 30 minutes at 37°C. The reaction mixture was then extracted with phenol-chloroform, precipitated with ethanol, denatured in formamide at 70°C, and subjected to electrophoresis on a 6 percent acrylamide, 8*M* urea, 1× TBE gel.

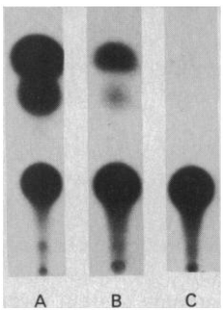


Fig. 9. Chloramphenicol acetyltransferase assay in JURKAT cells transfected with pSV2CAT (lane A), or the IL-2 receptor promoter construct in $pA_{10}CAT_2$ in correct (lane B) or reverse (lane C) orientation. JURKAT acute lymphocytic leukemia T cells were transfected with DNA's by DEAE dextran, and CAT assays performed as described (*33*). Cells (10^7) in the log phase of growth were washed twice with serum-free media, suspended in 1 ml of RPMI 1640 media containing DEAE dextran (250 μg/ml) and plasmid DNA (5 μg). The cells were incubated for 25 minutes at 37°C, cultures were then diluted to 10 ml in RPMI 1640 media containing 10 percent fetal bovine serum (FBS) and 100 μ*M* chloroquine, and incubated for an additional 30 minutes at 37°C. The cells were washed once and cultured in 10 percent FBS. After 24 hours, PMA (50 ng/ml) was added to the cultures. After an additional 24 hours of culture, cells were washed with phosphate-buffered saline, suspended in 1 ml of TEN (25 m*M* tris, *p*H 7.5, 1 m*M* EDTA, 17 m*M* NaCl), centrifuged, resuspended in 150 μl of 0.25*M* tris, *p*H 7.8, and frozen and thawed three times; they were then centrifuged at 12,000*g* for 10 minutes. The supernatant extract (150 μl) was combined with 20 μl of 12 m*M* acetyl coenzyme A (Pharmacia P-L) and 10 μl of [^{14}C]chloramphenicol (New England Nuclear) and incubated for 2 1/2 hours at 37°C. Reaction mixtures were extracted with ethyl acetate, dried, resuspended in 10 μl of ethyl acetate, separated by thin-layer chromatography, and autoradiographed. The percentage conversions of [^{14}C]chloramphenicol into acetylated chloramphenicol were as follows: Lane A, 54.4 percent; lane B, 4.0 percent; lane C, 0.28 percent. The promoter fragment used has its 3′ end in the sequence corresponding to "Promoter region plus exon 1" (see Fig. 1). It extends from the Pst I site (CTGCAG), 54 bp 5′ to the ATG translation initiation site to the next upstream Pst I site, approximately 1.4 kb further 5′.

end of the cDNA and a second 58 bases further 5′. In HUT-102B2 and four other HTLV-I infected ATL cells, but not in normal activated T cells, a third signal further 5′ was also detected. S1 nuclease studies confirmed that two major transcription initiation sites existed in the genomic DNA segment immediately 5′ to the first exon. Using the CAT assay, we have demonstrated that this region contains functional promoter activity. The third transcription initiation site present in the five ATL lines was located approximately 250 bases beyond the end of the cDNA by primer extension, but only 190 bases beyond in S1 nuclease assays. These data are most consistent with the third transcription initiation site being located 5′ to the fragment used in the S1 nuclease assay and that the transcript splices into the fragment of interest. At present, the precise location of this third putative initiation site and promoter have not been mapped.

It is known that IL-2 receptors are constitutively expressed at high levels in ATL cells, but must be induced in normal activated T cells. The apparent presence of a third promoter in ATL but not normal T cells is therefore intriguing; however, the strength of this promoter appears less than that of the other two promoters. Although it is conceivable that transcripts from this third promoter are more stable or are translated more efficiently than those of the other two, we doubt that this third promoter fully accounts for the large number of receptors continuously expressed on ATL cells.

Comparison of the IL-2 receptor promoter region to sequences available in the NBRF and GenBank databases have not revealed any significant homologies. Specifically the IL-2 receptor promoter region does not share homology to the IL-2 promoter.

Multiple mRNA species originating from a single gene, as is the case with the IL-2 receptor, have been described in other systems. For example, c-*myc* has two transcription initiation sites (*27*). Many examples of genes that manifest alternative mRNA splicing or differential polyadenylation (or both) have also been noted (*28*). The IL-2 receptor combines all three of these processes. Coupling the data we provide here for three different transcription initiation sites to earlier data for alternative mRNA splicing and at least three different polyadenylation signals, we can conclude that perhaps as many as 18 different mature mRNA forms can be produced by this single gene. The events that regulate which of these mRNA's are made and when remain to be elucidated.

The second and fourth exons of IL-2 receptor share significant homology with human complement factor B, specifically the Ba fragment which serves a recognition function. It is noteworthy that two membrane receptors, the IL-2 receptor and LDL receptor, each have unexpected homology to proteins involved in proteolytic cascades: complement and coagulation factors, respectively. Further, the IL-2 receptor contains an internal gene duplication involving exons 2 and 4. The stoichiometry of and specific site or sites of IL-2 binding to IL-2 receptors are unknown, but the internal gene duplication raises the interesting possibility that the duplicated domains may be critical to IL-2 binding and that there may be two IL-2 binding sites per receptor molecule.

References and Notes

1. D. A. Morgan, F. W. Ruscetti, R. C. Gallo, *Science* **193**, 1007 (1976).
2. R. J. Robb, A. Munck, K. A. Smith, *J. Exp. Med.* **154**, 1455 (1981).
3. A. Coutinho, E. L. Larsson, K. O. Grönvik, J. Andersson, *Eur. J. Immunol.* **9**, 587 (1979); S. Gillis, P. E. Baker, F. W. Ruscetti, K. A. Smith, *J. Exp. Med.* **148**, 1093 (1978).
4. M. Tsudo, T. Uchiyama, H. Uchino, *J. Exp. Med.* **160**, 612 (1984).
5. T. A. Waldman *et al.*, *ibid.*, p. 1450.
6. W. J. Leonard *et al.*, *Nature (London)* **300**, 267 (1982); W. J. Leonard, J. M. Depper, R. J. Robb, T. A. Waldmann, W. C. Greene, *Proc. Natl. Acad. Sci. U. S. A.* **80**, 6957 (1983).
7. Y. Wano *et al.*, *J. Immunol.* **132**, 3005 (1984).
8. W. J. Leonard, J. M. Depper, M. Krönke, R. J. Robb, T. A. Waldmann, W. C. Greene, *J. Biol. Chem.* **260**, 1872 (1985).
9. D. A. Shackleford and I. S. Trowbridge, *ibid.* **259**, 11706 (1984).
10. J. M. Depper, W. J. Leonard, M. Krönke, T. A. Waldmann, W. C. Greene, *J. Immunol.* **133**, 1691 (1984).
11. W. J. Leonard *et al.*, *Nature (London)* **312**, 626 (1984).
12. T. Nikaido *et al.*, *ibid.*, p. 631.
13. D. Cosman *et al.*, *ibid.*, p. 768.
14. W. J. Leonard *et al.*, *J. Cell. Biochem.*, in press.
15. W. J. Leonard, T. A. Donlon, R. V. Lebo, W. C. Greene, *Science* **228**, 1547 (1985).
16. M. Krönke, W. J. Leonard, J. M. Depper, W. C. Greene, *ibid.*, p. 1215.
17. The libraries screened were a fetal liver library prepared in the laboratory of T. Maniatis and a peripheral blood leukocyte library prepared in the laboratory of P. Leder.
18. C. L. Queen and L. J. Korn, *Methods Enzymol.* **65**, 595 (1980).
19. J. Miller *et. al.*, *J. Immunol.* **134**, 4212 (1985).
20. D. W. Russell *et al.*, *Cell* **37**, 577 (1984).
21. T. C. Südhof, J. L. Goldstein, M. S. Brown, D. W. Russell, *Science* **228**, 815 (1985).
22. W. J. Leonard, M. Krönke, N. J. Peffer, J. M. Depper, W. C. Greene, *Proc. Natl. Acad. Sci. U. S. A.* **82**, 6281 (1985).
23. W. C. Greene, R. J. Robb, P. B. Svetlik, J. M. Depper, W. J. Leonard, *J. Exp. Med.* **162**, 363 (1985).
24. This program was written by Dr. Minoru Kanehisa, Kyoto University.
25. J. E. Mole, J. K. Anderson, E. A. Davison, D. E. Woods, *J. Biol. Chem.* **259**, 3407 (1984).
26. J. Miller and R. Germain, personal communication.
27. J. Battey, C. Moulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, P. Leder, *Cell* **34**, 779 (1983).
28. J. E. Darnell, Jr., *Nature (London)* **297**, 365 (1982).
29. F. Sanger and A. R. Coulson, *J. Mol. Biol.* **94**, 414 (1975); G. R. Crabtree, C. Comeau, D. Fowlkes, J. A. Kant, *ibid.*, in press.
30. M. B. Shapiro and P. Senapathy, *Nucleic Acids Res.*, in press.
31. J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
32. M. Krönke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5214 (1984).
33. C. M. Gorman, F. Moffat, B. H. Howard, *Mol. Cell. Biol.* **2**, 1044 (1982); C. M. Gorman, G. T. Merlino, M. C. Willingham, I. Pastan, B. H. Howard, *Proc. Natl. Acad. Sci. U. S. A.* **79**, 6777 (1982).

34. We thank N. Holbrook, National Cancer Institute, for assistance with performing the chloramphenicol acetyltransferase assays; M. Shapiro, Division of Computer Research and Technology, NIH, for assistance with preparation of the DNA sequence for publication; R. M. Belagaje, Lilly Research Laboratories, for synthesis of the oligonucleotide used in the primer extension studies; J. C. Jenson, Hoffmann–La Roche, Inc., for performing several computer analyses; M. Popovic, National Cancer Institute, for ATL cell lines; and G. R. Crabtree, Stanford University, for critical comments.

31 July 1985; accepted 25 September 1985