Volume 12 Number 12 1984 Nucleic Acids Research

DNA sequence of the 5' flanking region of the human interleukin 2 gene: homologies with adult T-cell leukemia virus

Nikki J.Holbrook, Michael Lieber and Gerald R.Crabtree

Laboratory of Pathology, National Cancer Institute, Building 10, Room 2N-113, Bethesda, MD 20205, USA

Received 14 May 1984; Accepted 23 May 1984

ABSTRACT

We previously reported the sequence of the gene for human interleukin 2 (IL-2) which included about 220 b.p. of the 5' flanking region. In this report we have extended the 5' sequence to 1263 b.p. before the start site of transcription and have compared the entire sequence of the IL-2 gene and its 5' flanking region with that of Adult T-cell leukemia virus (ATLV). Four regions of limited homology were noted. Three areas of homology to the viral long terminal repeat (LTR) were seen in the 5' sequence flanking IL-2. A fourth region of homology to the viral LTR was located within the second intron of the IL-2 gene. While the significance of these homologies is not known, it is possible that they are involved in regulating the expression of ATLV and IL-2, both of which are generally restricted to T cells.

INTRODUCTION

Interleukin 2 (IL-2), also known as T-cell growth factor (TCGF), is a lymphokine produced by lectin- or antigen-stimulated T cells which is capable of initiating and maintaining long term in vitro growth of activated T cells (1). Although the protein has been purified to homogeneity (2,3) and its biological effects studied extensively (2-5), little is known about its regulation at a genetic level. In addition to antigen, mitogens such as phytohemagglutinin and concanavalin A induce IL-2 production in human T lymphocytes (2,6). The tumor promoter phorbol myristic acetate, by itself does not induce IL-2 production, but enhances the induction by mitogen or antigen (2, 6). Recently, a monoclonal antibody to a cell surface protein on T cells (T3) was shown to induce IL-2 mRNA (7). This latter effect is probably mediated by the same mechanism which antigens use to activate cell division, since T3 appears to be part of the T cell antigen receptor complex. Thus it is likely that the activation of the IL-2 gene will serve as a useful model for studying the way in which antigens initiate the orderly and sequential activation of genes that lead to clonal expansion of T lymphocytes.

To study the processes involved in regulation of IL-2 at the molecular level we cloned the gene from a human DNA library. Recently, we (8) and

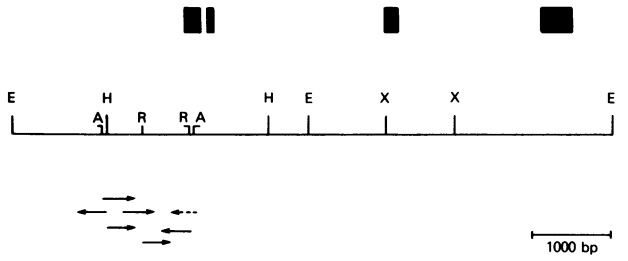

Figure 1. Restriction map and sequencing strategy for the 5' flanking region of the human IL-2 gene. Rectangles indicate the exons of the gene. Arrows indicate the direction and extent of sequence analysis of individual fragments. The sequence is represented 5' to 3' from left to right. E, EcoR I; H, Hind III; X, Xba I; A, Alu I; R, Rsa I. The dotted arrow indicates the overlapping 5' end of our previously published sequence.

Fugita et al. (9) independently reported the structure and complete nucleotide sequence of the IL-2 gene. The gene, which is comprised of four exons separated by intervening sequences, occurs as a single copy in the genome. Few discrepancies occur in the two reported sequences. Most of these involve single base-pair substitutions or deletions in the third intron of the gene. At present, the published nucleic acid sequence for the 5' flanking region of the IL-2 gene extends about 430 b.p. upstream from the site for initiation of transcription (9). In view of the potential role of DNA sequences farther 5' influencing expression of IL-2 we sequenced an additional 900 b.p. of the 5' flanking region. Comparision of the sequence of IL-2 with the sequence for Adult T-cell leukemia virus (ATLV) (10,11) revealed several areas of homology within the 5' end of IL-2 and the long terminal repeat (LTR) of ATLV.

MATERIALS AND METHODS

Materials. DNA polymerase was from Boehringer Mannheim; restriction endonucleases were from New England BioLabs; T4 DNA ligase, M13 phage vectors and unlabeled deoxynucleotide triphosphates and dideoxynucleotide triphosphates were from P-L Biochemicals; and [$-^{32}$P]deoxynucleotide triphosphates were from Amersham.

DNA sequencing and analysis. The DNA sequences were determined by the chain termination method of Sanger et al. (12) as described previously (8). Sequences were analyzed and compared on an IBM system 370 computer using a program described by Queen and Korn (13) and were analyzed graphically by standard dot matrix plots (14).

```
                                                                      60
AGTGGTTTTT  GGAGTCAGTA  CATTCTCTTT  TCAAATCCTT  CTCTGCCCCT  TACTGGCAAT
                                                                     120
AAGGGCTGAG  TGACCTAGAG  GCAAATTACT  TAACTTCTCT  GAGCCTCAGT  TTTCTAATCT
                                                                     180
GCAAAATAGG  AGCCATCACT  TCACAAGTCT  GTAAGACTTA  TATTAGACTA  AGTGCCTGCC
                                                                     240
TGTACACTGT  TCTCTTTTTC  TCTCTTTCTA  TATACCTGAA  GGCATTATAG  TGCTAGATGT
                                                                     300
CTGTTTAAAG  ACCAGACAAT  ATTGTCTTAA  AAAAACAAAC  AAAAACACAG  ACAATACCAT
                                                                     360
CTTTAAAAAA  AAAAAAAAAG  TCCAGGTAAG  AAATAAATAA  GGCCATAGAA  TGGAAGCTTT
                                                                     420
ACAAGGACTC  TCTTTGAGAC  AGGATCTCCT  CAAGTGTCCC  CAGGTTAAAT  TAGAAGTATA
                                                                     480
TATCCGTACA  ATTGTTCAGC  CAGTTTGTGC  ACTGTACTGA  GGATGAATGA  ACACCTATCC
                                                                     540
TAAATATCCT  AGTCTTCTGA  CTAAAAACAA  GATCATATTT  CATAACGATT  ATTGTTACAT
                                                                     600
TCATAGTGTC  CCAGGTGATT  TAGAGGATAA  ATAAAAATCC  ATTAAAGAGG  TAAAGACATA
                                                                     660
AAAACGAGAA  ACATGGACTG  GTTTACACAT  AACACATACA  AAGTCTATTA  TAAAACTAGC
                                                                     720
ATCAGTATCC  TTGAATCGAA  ACCTTTTTCT  GAGTATTTAA  CAATCGCACC  CTTTAAAAAA
                                                                     780
TGTACATAGA  CATTAAGAGA  CTTAAACAGA  TATATAATCA  TTTTAAATTA  AAATAGCGTT
                                                                     840
AAACAGTACC  TCAAGCTCAA  TAAGCATTTT  AAGTATTCTA  ATCTTAGTAT  TTCTCTAGCT
                                                                     900
GACATGTAAG  AAGCAATCTA  TCTTATTGTA  TGCAATTAGC  TCATTGTGTG  GATAAAAAGG
                                                                     960
TAAAACCATT  CTGAAACAGG  AAACCAATAC  ACTTCCTGTT  TAATCAACAA  ATCTAAACAT
                                                                    1020
TTATTCTTTT  CATCTGTTTA  CTCTTGCTCT  TGTCCACCAC  AATATGCTAT  TCACATGTTC
                                                                    1080
AGTGTAGTTT  TATGACAAAG  AAAATTTTCT  GAGTTACTTT  TGTATCCCCA  CCCCCTTAAA
                                                                    1140
GAAAGGAGGA  AAAACTGTTT  CATACAGAAG  GCGTTAATTG  CATGAATTAG  AGCTATCACC
                                                                    1200
TAAGTGTGGG  CTAATGTAAC  AAAGAGGGAT  TTCACCTACA  TCCATTCAGT  CAGTCTTTGG
                                                                    1260
GGGTTTAAAG  AAATTCCAAA  GAGTCATCAG  AAGAGGAAAA  ATGAAGGTAA  TGTTTTTTCA
                                                                    1320
GACTGGTAAA  GTCTTTGAAA  ATATGTGTAA  TATGTAAAAC  ATTTTGACAC  CCCCATAATA
                                                                    1380
TTTTTCCAGA  ATTAACAGTA  TAAATTGCAT  CTCTTGTTCA  AGAGTTCCCT  ATCACTCTTT
                                                                    1433

AATCACTACT  CACAGTAACC  TCAACTCCTG  CCACA
                                        ATG TAC AGG ATG CAA CTC
                                        Met Tyr Arg Met Gln Leu
                                                                    1490
CTG TCT TGC ATT GCA CTA AGT CTT GCA CTT GTC ACA AAC AGT GCA CCT ACT TCA AGT
Leu Ser Cys Ile Ala Leu Ser Leu Ala Leu Val Thr Asn Ser Ala Pro Thr Ser Ser
```

Figure 2.

```
                                                                    1547
TCT ACA AAG AAA ACA CAG CTA CAA CTG GAG CAT TTA CTG CTG GAT TTA CAG ATG ATT
Ser Thr Lys Lys Thr Gln Leu Gln Leu Glu His Leu Leu Leu Asp Leu Gln Met Ile

                                                                    1602
TTG AAT GGA ATT AAT        GTAAGTATAT   TTCCTTTCTT   ACTAAAATTA   TTACATTTAG
Leu Asn Gly Ile Asn

                                                                    1661
TAATCTAGCT   GGAGATCATT   TCTTAATAAC   AATGCATTAT   ACTTTCTTAG AAT TAC AAG
                                                                Asn Tyr Lys

                                                                    1712
AAT CCC AAA CTC ACC AGG ATG CTC ACA TTT AAG TTT TAC ATG CCC AAG AAG
Asn Pro Lys Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys Lys

                                                                    1772
GTAAGTACAA   TATTTTATGT   TCAATTTCTG   TTTTAATAAA   ATTCAAAGTA   ATATGAAAAT

                                                                    1832
TTGCACAGAT   GGGACTAATA   GCAGCTCATC   TGAGGTAAAG   AGTAACTTTA   ATTTGTTTTT

                                                                    1892
TTGAAAACCC   AAGTTTGATA   ATGAAGCCTC   TATTAAAACA   GTTTTACCTA   TATTTTTAAT

                                                                    1952
ATATATTTGT   GTGTTGGTGG   GGGTGGGAGA   AAACATAAAA   ATAATATTCT   CTCACTTTAT

                                                                    2012
CGATAAGACA   ATTCTAAACA   AAAATGTTCA   TTTATGGTTT   CATTTAAAAA   TGTAAAACTC

                                                                    2072
TAAAATATTT   GATTATGTCA   TTTTAGTATG   TAAAATACCA   AAATCTATTT   CCAAGGAGCC

                                                                    2132
CACTTTTAAA   AATCTTTTCT   TGTTTTAGGA   AAGGTTTCTA   AGTGAGAGGC   AGCATAACAC

                                                                    2192
TAATAGCACA   GAGTCTGGGG   CCAGATATCT   GAAGTGAAAT   CTCAGCTCTG   CCATGTCCTA

                                                                    2252
GCTTTCATGA   TCTTTGGCAA   ATTACCTACT   CTGTTTGTGA   TTCAGTTTCA   TGTCTACTTA

                                                                    2312
AATGAATAAC   TGTATATACT   TAATATGGCT   TTGTGAGAAT   TAGTAAGTAA   ATGTAAAGCA

                                                                    2372
CTCAGAACCG   TGTCTGGCAT   AAGGTAAATA   CCATACAAGC   ATTAGCTATT   ATTAGTAGTA

                                                                    2432
TTAAAGATAA   AATTTTCACT   GAGAAATACA   AAGTAAAATT   TTGGACTTTA   TCTTTTTACC

                                                                    2492
AATAGAACTT   GAGATTTATA   ATGCTATATG   ACTTATTTTC   CAAGATTAAA   AGCTTCATTA

                                                                    2552
GGTTGTTTTT   GGATTCAGAT   AGAGCATAAG   CATAATCATC   CAAGCTCCTA   GGCTACATTA

                                                                    2612
GGTGTGTAAA   GCTACCTAGT   AGCTGTGCCA   GTTAAGAGAG   AATGAACAAA   ATCTGGTGCC

                                                                    2672
AGAAAGAGCT   TGTGCCAGGG   TGAATCCAAG   CCCAGAAAAT   AATAGGATTT   AAGGGGACAC

                                                                    2732
AGATGCAATC   CCATTGACTC   AAATTCTATT   AATTCAAGAC   AAATCTGCTT   CTAACTACCC

                                                                    2792
TTCTGAAAGA   TGTAAAGGAG   ACAGCTTACA   GATGTTACTC   TAGTTTAATC   AGAGCCACAT

                                                                    2862
AATGCAACTC   CAGCAACATA   AAGATACTAG   ATGCTGTTTT   CTGAAGAAAA   TTTCTCCACA

                                                                    2912
TTGTTCATGC   CAAAAACTTA   AACCCGAATT   TGTAGAATTT   GTAGTGGTGA   ATTGAAAGCG

                                                                    2972
CAATAGATGG   ACATATCAGG   GGATTGGTAT   TGTCTTGACC   TACCTTTCCC   ACTAAAGAGT

                                                                    3032
GTTAGAAAGA   TGAGATTATG   TGCATAATTT   AGGGGTGGTA   GAATTCATGG   AAATCTAAGT
```

```
                                                                    3092
TTGAAACCAA  AAGTAATGAT  AAACTCTATT  CATTTGTTCA  TTTAACCCTC  ATTGCACATT

                                                                    3152
TACAAAAGAT  TTTAGAAACT  AATAAAAATA  TTTGATTCCA  AGGATGCTAT  GTTAATGCTA

                                                                    3212
TAATGAGAAA  GAAATGAAAT  CTAATTCTGG  CTCTACCTAC  TTATGTGGTC  AAATTCTGAG

                                                                    3272
ATTTAGTGTG  CTTATTTATA  AAGTGGAGAT  GATACTTCAC  TGCCTACTTC  AAAAGATGAC

                                                                    3332
TGTGAGAAGT  AAATGGGCCT  ATTTTGGAGA  AAATTCTTTT  AAATTGTAAT  ATACCATAGA

                                                                    3392
AATATGAAAT  ATTATATATA  ATATAGAATC  AAGAGGCCTG  TCCAAAAGTC  CTCCCAAAGT

                                                                    3452
ATTATAATCT  TTTATTTCAC  TGGGACAAAC  ATTTTTAAAA  TGCATCTTAA  TGTAGTGATT

                                                                    3512
GTAGAAAAGT  AAAAATTTAA  GACATATTTA  AAAATGTGTC  TTGCTCAAGG  CTATATTGAG

                                                                    3572
AGCCACTACT  ACATGATTAT  TGTTACCTAG  TGTAAAATGT  TGGGATTGTG  ATAGATGGCA

                                                                    3632
TCCAAGAGTT  CCTTCTCTCT  CAACATTCTG  TGATTCTTAA  CTCTTAGACT  ATCAAATATT

                                                                    3692
ATAATCATAG  AATGTGATTT  TTATGCCTTC  CACATTCTAA  TCTCATCTGG  TTCTAATGAT

                                                                    3752
TTTCTATGCA  GATTGGAAAA  GTAATCAGCC  TACATCTGTA  ATAGGCATTT  AGATGCAGAA

                                                                    3812
AGTCTAACAT  TTTGCAAAGC  CAAATTAAGC  TAAAACCAGT  GAGTCAACTA  TCACTTAACG

                                                                    3872
CTAGTCATAG  GTACTTGAGC  CCTAGTTTTT  CCAGTTTTAT  AATGTAAACT  CTACTGGTCC

                                                                    3932
ATCTTTACAG  TGACATTGAG  AACAGAGAGA  ATGGTAAAAA  CTACATACTG  CTACTCCAAA

                                                                    3992
TAAAATAAAT  TGGAAATTAA  TTTCTGATTC  TGACCTCTAT  GTAAACTGAG  CTGATGATAA

                                                                    4049
TTATTATTCT  AG  GCC ACA GAA CTG AAA CAT CTT CAG TGT CTA GAA GAA GAA CTC AAA
                    Ala Thr Glu Leu Lys His Leu Gln Cys Leu Glu Glu Glu Leu Lys

                                                                    4106
CCT CTG GAG GAA GTG CTA AAT TTA GCT CAA AGC AAA AAC TTT CAC TTA AGA CCC AGG
Pro Leu Glu Glu Val Leu Asn Leu Ala Gln Ser Lys Asn Phe His Leu Arg Pro Arg

                                                                    4158
GAC TTA ATC AGC AAT ATC AAC GTA ATA GTT CTG GAA CTA AAG          GTAAGGCATT
Asp Leu Ile Ser Asn Ile Asn Val Ile Val Leu Glu Leu Lys

                                                                    4218
ACTTTATTTG  CTCTCCTGGA  AATAAAAAAA  AAAAAGTAGG  GGGAAAAGTA  CCACATTTTA

                                                                    4278
AAGTGACATA  ACATTTTTGG  TATTTGTAAA  GTACCCATGC  ATGTAATTAG  CCTACATTTT

                                                                    4338
AAGTACACTG  TGAACATGAA  TCATTTCTAA  TGTTAAATGA  TTAACTGGGG  AGTATAAGCT

                                                                    4398
ACTGAGTTTG  CACCTACCAT  CTACTAATGG  ACAAGCCTCA  TCCCAAACTC  CATCACCTTT

                                                                    4458
CATATTAACA  CAAAACTGGG  AGTGAGAGAG  AAGTGACTGA  GTTGAGTTTC  ACAGAAACGC

                                                                    4518
AGGCAAGATT  TTATTATATA  TTTTTCAAGT  TCCTTCACAG  ATCATTTACT  GGAATAGCCA

                                                                    4578
ATACTGAGTT  ACCTGAAAGG  CTTTTCAAAT  GGTGTTTCCT  TATCATTTGA  TGGAAGGACT

                                                                    4638
ACCCATAAGA  GATTTGTCTT  AAAAAAAAAA  ACTGGAGCCA  TTAAAATGGC  CAGTGGACTA
```

```
                                                                    4698
AACAAACAAC  AATCTTTTTA  GAGGCAATCC  CACTTTCAGA  ATCTTAAGTA  TTTTTAAATG

                                                                    4758
CACAGGAAGC  ATAAAATATG  CAAGGGACTC  AGGTGATGTA  AAAGAGATTC  ACTTTTGTCT

                                                                    4818
TTTTATATCC  CGTCTCCTAA  GGTATAAAAT  TCATGAGTTA  ATAGGTATCC  TAAATAAGCA

                                                                    4878
GCATAAGTAT  AGTAGTAAAA  GACATTCCTA  AAAGTAACTC  CAGTTGTGTC  CAAATGAATC

                                                                    4938
ACTTATTAGT  GGACTGTTTC  AGTTGAATTA  AAAAAATACA  TTGAGATCAA  TGTCATCTAG

                                                                    4998
ACATTGACAG  ATTCAGTTCC  TTATCTATGG  CAAGAGTTTT  ACTCTAAAAT  AATTAACATC

                                                                    5058
AGAAAACTCA  TTCTTAACTC  TTGATACAAA  TTTAAGACAA  AACCATGCAA  AAATCTGAAA

                                                                    5118
ACTGTGTTTC  AAAAGCCAAA  CACTTTTTAA  AATAAAAAAA  TCCCAAGATA  TGACAATATT

                                                                    5178
TAAACAATTA  TGCTTAAGAG  GATACAGAAC  ACTGCAACAG  TTTTTTAAAA  GAGAATACTT

                                                                    5238
ATTTAAAGGG  AACACTCTAT  CTCACCTGCT  TTTGTTCCCA  GGGTAGGAAT  CACTTCAAAT

                                                                    5298
TTGAAAAGCT  CTCTTTTAAA  TCTCACTATA  TATCAAAATA  GTTGCCTCCT  TAGCTTATCA

                                                                    5358
ACTAGAGGAA  GCGTTTAAAT  AGCTCCTTTC  AGCAGAGAAG  CCTAATTTCT  AAAAAGCCAG

                                                                    5418
TCCACAGAAC  AAAATTTCTA  ATGTTTAAAG  CTTTTAAAAG  TTGGCAAATT  CACCTGCATT

                                                                    5478
GATACTATGA  TGGGGTAGGG  ATAGGTGTAA  GTATTTATGA  AGATGTTCAT  TCACACAAAT

                                                                    5538
TTACCCAAAC  AGGAAGCATG  TCCTACCTAG  CTTACTCTAG  TGTAGCTCGT  TTCGTCTTTG

                                                                    5598
GGGAAAATAT  AAGGAGATTC  ACTTAAGTAG  AAAAATAGGA  GACTCTAATC  AAGATTTAGA

                                                                    5658
AAAGAAGAAA  GTATAATGTG  CATATCAATT  CATACATTTA  ACTTACACAA  ATATAGGTGT

                                                                    5718
ACATTCAGAG  GAAAAGCGAT  CAAGTTTATT  TCACATCCAG  CATTTAATAT  TTGTCTAGAT

                                                                    5778
CTATTTTTAT  TTAAATCTTT  ATTTGCACCC  AATTTAGGGA  AAAAATTTTT  GTGTTCATTG

                                                                    5838
ACTGAATTAA  CAAATGAGGA  AAATCTCAGC  TTCTGTGTTA  CTATCATTTG  GTATCATAAC

                                                                    5898
AAAATACGCA  ATTTTGGCAT  TCATTTTGAT  CATTTCAAGA  AAATGTGAAT  AATTAATATG

                                                                    5958
TTTGGTAAGC  TTGAAAATAA  AGGCAACAGG  CCTATAAGAC  TTCAATTGGG  AATAACTGTA

                                                                    6015
TATAAGGTAA  ACTACTCTGT  ACTTTAAAAA  ATTAACATTT  TTCTTTTATA  G GGA TCT
                                                              Gly Ser

                                                                    6072
GAA ACA ACA TTC ATG TGT GAA TAT GCT GAT GAG ACA GCA ACC ATT GTA GAA TTT CTG
Glu Thr Thr Phe Met Cys Glu Tyr Ala Asp Glu Thr Ala Thr Ile Val Glu Phe Leu

                                                                    6127
AAC AGA TGG ATT ACC TTT TGT CAA AGC ATC ATC TCA ACA CTG ACT      TGATAATTAA
Asn Arg Trp Ile Thr Phe Cys Gln Ser Ile Ile Ser Thr Leu Thr

                                                                    6187
GTGCTTCCCA  CTTAAAACAT  ATCAGGCCTT  CTATTTATTT  AAATATTTAA  ATTTTATATT
```

```
                                                                    6247
TATTGTTGAA  TGTATGGTTT  GCTACCTATT  GTAACTATTA  TTCTTAATCT  TAAAACTATA

                                                                    6307
AATATGGATC  TTTTATGATT  CTTTTTGTAA  GCCCTAGGGG  CTCTAAAATG  GTTTCACTTA

                                                                    6367
TTTATCCCAA  AATATTTATT  ATTATGTTGA  ATGTTAAATA  TAGTATCTAT  GTAGATTGGT

                                                                    6427
TAGTAAAACT  ATTTAATAAA  TTTGATAAAT  ATAAACAAGC  CTGGATATTT  GTTATTTTGG

                                                                    6487
AAACAGCACA  GAGTAAGCAT  TTAAATATTT  CTTAGTTACT  TGTGTGAACT  GTAGGATGGT

                                                                    6547
TAAAATGCTT  ACAAAAGTCA  CTCTTTCTCT  GAAGAAATAT  GTAGAACAGA  GATGTAGACT

                                                                    6607
TCTCAAAAGC  CCTTGCTTTG  TCCTTTCAAG  GGCTGATCAG  ACCCTTAGTT  CTGGCATCTC

                                                                    6667
TTAGCAGATT  ATATTTTCCT  TCTTCTTAAA  ATGCCAAACA  CAAACACTCT  TGAAACTCTT

CATAGATTTG  GTGTGGC
```

Figure 2. Complete DNA sequence of the human IL-2 gene and its 5' flanking region. The amino acid sequences encoded by exons of IL-2 are shown below the DNA sequences. The 3' and 5' untranslated regions are underlined. The TATA box, putative transcription initiation site, translation initiation site, termination codon and polyadenylation signal are enclosed in rectangles.

RESULTS AND DISCUSSION

The strategy for sequencing the 5' region flanking the IL-2 gene is shown in Fig. 1. The sequence obtained is shown in Fig. 2. Nucleotides 1-1133 represent newly determined sequence. For clarity and completeness we have also included our previously published sequence of the IL-2 gene.

Like the 3' untranslated region of IL-2, the 5' flanking DNA is highly rich in A residues. In particular, the segment extending from nucleotides 268-340 is comprised of 67% A residues with a long stretch of 14 A's and shares considerable homology with Alu family sequences (15).

Expression of Adult T-cell leukemia virus (ATLV), like that of IL-2 is highly specific for T cells. Using the computer program of Korn and Queen as well as dot matrices analyses we have compared the entire IL-2 gene and its 5' flanking sequences with those of ATLV (11,16). Four regions of limited homology with the viral LTR were noted, three of which were present in the 5' region of IL-2. These areas are depicted in Fig. 3. It is of interest to note that the homologous sequences shown in 3B occur at similar distances from the transcription initiation sites of the viral and IL-2 genes. For the virus this encompasses the cellular flanking sequence and 5' terminus of the LTR in one published sequence (11). Another published sequence of ATLV is

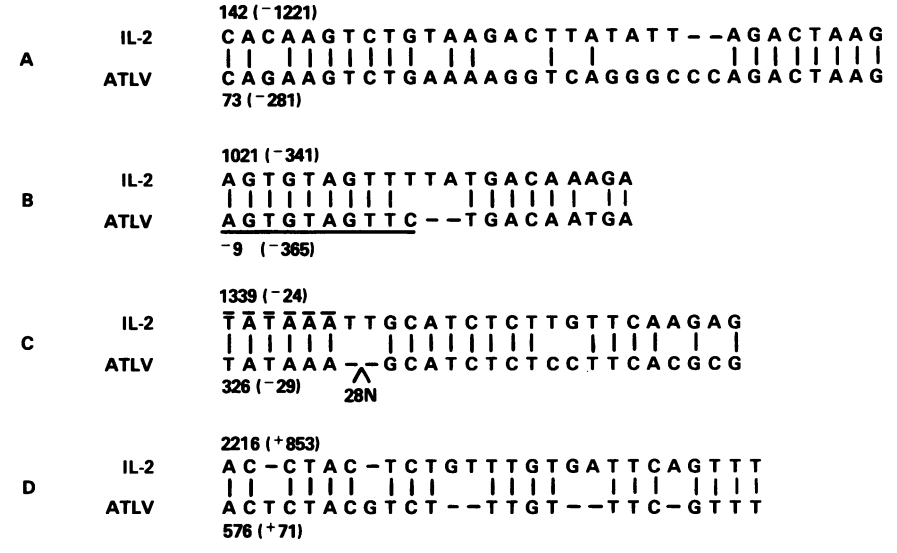

Figure 3. Homologies between ATLV and the IL-2 gene. Numbers without parentheses refer to the position relative to the first position in the published sequence. The numbers in parentheses give the relative position of the nucleotide to the transcription initiation site.

flanked by cellular sequences which have little or no homology to the IL-2 gene (16). The region of homology depicted in 3C extends from the TATA box up to the transcription initiation site for IL-2 and is associated with the R sequence, or terminal repeat sequence (16), within the LTR of ATLV. A fourth region of homology was noted within the second intron of the IL-2 gene and the R sequence of ATLV LTR.

Fugita et al. (9) noted two areas of the 5' sequence of the IL-2 gene which showed homology to the human gamma interferon gene sequence just 5' of the TATA box. We have found no other homologies within the additional 5' sequence of IL-2. Furthermore, the sequences of homology between IL-2 and human gamma interferon do not correspond to those regions of homology between IL-2 and ATLV.

The possible significance of the homologies between IL-2 and ATLV is at present unclear. However, since the expression of IL-2 and ATLV display specificity for T cells, it is possible that these regions are associated with the conferred specificity of expression.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Jacob B. Maizel and John Owens for their help with analysis by dot matrix plots. N.J.H. was supported by a fellowship from the Leukemia Society of America and a grant from the Eli Lilly Research Laboratories.

REFERENCES

1. Morgan, D.A., Ruscetti, F.W. and Gallo, R.C. (1976) Science 193, 1007-1008
2. Robb, R.J. and Smith, K.A. (1981) Mol. Immunol. 18, 1087-1094
3. Robb, R.J., Munck, A. and Smith, K.A. (1981) J. Exp. Med. 154, 1455-1474
4. Smith, K.A. (1980) Immunol. Rev. 51, 337-357
5. Rosenberg, S.A., Grimm, E.A., McGrogan, M., Doyle, M., Kawasaki, E., Koths, K. and Mark, D.F. (1984) Science 223, 1412-1415
6. Gillis, S. and Watson, J. (1980) J. Exp. Med. 152, 1709-1719
7. Stobo, J.D., Wiskocil, R.L., Imboden, J.B., Kamin-Lewis, R.M. and Weiss, A. (1984) Federation Proceedings 43, 1734
8. Holbrook, N.J., Smith, K.A., Fornace, A.J., Comeau, C.M., Wiskocil, R. C. and Crabtree, G.R. (1984) Proc. Natl. Acad. Sci. USA 81, 1634-1638
9. Fujita, T., Takaoka, C. Matsui, H. and Taniguchi, T. (1983) Proc. Natl. Acad. Sci. USA 80, 7437-7441
10. Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. and Gallo, R.C. 1980 Proc. Natl. Acad. Sci. USA 77, 7415-7419
11. Seiki, M. Hattori, S., Hirayama, Y. and Yoshida, M. (1983) Proc. Natl. Acad. Sci. USA 80, 3618-3622
12. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467
13. Queen, C.L. and Korn, L.J. (1980) Methods Enzymol. 65, 595-609
14. Maizel, J.V. and Lenk, R.P. (1981) Proc. Natl. Acad. Sci. USA 78, 7665-7669
15. Rubin, C.M., Houck, C.M., Deininger, P.L., Friedmann, T. and Schmid, C. W. (1980) Nature 284, 372-374
16. Seiki, M., Hattori, S. and Yoshida, M. (1982) Proc. Natl. Acad. Sci. USA 79, 6899-6902