

---

**Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localization**

---

Glenn E.Nedwin\*, Susan L.Naylor+, Alan Y.Sakaguchi+, Douglas Smith, Julie Jarrett-Nedwin\*, Diane Pennica, David V.Goeddel and Patrick W.Gray

---

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080, and +Department of Cellular and Structural Biology, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78284, USA

---

Received 13 May 1985; Revised and Accepted 10 August 1985

---

**ABSTRACT**

Human Tumor Necrosis Factor and Lymphotoxin are cytotoxic proteins which have similar biological activities and share 30 percent amino acid homology. The single copy genes which encode these proteins share several structural features: each gene is approximately three kilobase pairs in length and is interrupted by three introns. In addition, these genes are closely linked and have been mapped to human chromosome 6. However, only the last exons of both genes, which code for more than 80 percent of each secreted protein, are significantly homologous (56 percent).

**INTRODUCTION**

Tumor necrosis factor (TNF- $\alpha$ ) and Lymphotoxin (TNF- $\beta$ ) are two cytotoxic proteins secreted from mitogen activated peripheral blood leukocytes (PBLs) and cell lines of hematopoietic origin. TNF- $\alpha$  and - $\beta$  have very similar biological activities. Both agents cause the direct cytolysis of certain tumor cell lines in vitro and have antiproliferative activity on other tumor cell lines (1-8). TNF- $\alpha$  and - $\beta$  also act synergistically with interferon-gamma (IFN- $\gamma$ ) to kill various transformed cell lines (8-12). The growth of primary cells and normal cell lines, however, is not inhibited by these factors. Both TNFs cause the hemorrhagic necrosis of the murine Meth A sarcoma in vivo (3,4,8,13,14). These studies suggest that TNFs may be useful antitumor agents in the treatment of cancer.

We have recently cloned and expressed in E. coli the cDNAs for human TNF- $\beta$  (13) and human TNF- $\alpha$  (14). TNF- $\alpha$  and TNF- $\beta$  are each encoded by unique genes and share approximately 30 percent amino acid homology (13,14). TNF- $\beta$  is secreted from induced T-lymphocytes and is a glycosylated protein of 171 amino acids (monomer  $M_r$  25,000) (15,16). TNF- $\alpha$  is derived from activated monocytes and has a size of 157 residues (monomer  $M_r$  17,000) (14,17). These cytotoxic agents exhibit little species specificity in their anticellular activities (18-20). They appear to be the major cytolytic

factors produced by PBLs which have activity on the murine L-929 cell assay (10,13,14). Although the biological activities of TNF- $\alpha$  and TNF- $\beta$  are very similar, they are derived from different cell types and have distinct induction kinetics: TNF- $\beta$  is secreted from T-lymphocytes 24-48 hours following induction, while TNF- $\alpha$  is secreted from monocytes 4-24 hours after induction (21,22). Because these genes encode proteins of similar structure and activity but their gene expression is distinct, we were interested in analyzing and comparing their genomic structure.

### MATERIALS AND METHODS

The human TNF- $\beta$  gene was isolated from a recombinant human- $\lambda$  library (courtesy of Dr. William Wood, 23). Approximately  $7.5 \times 10^5$  plaques were screened by hybridizing (24) a  $^{32}\text{P}$ -labeled DNA probe encoding the 34 amino-terminal residues of TNF- $\beta$  (13). Filters were washed twice at low stringency (25,26) using 1X SSC (0.15 M NaCl, 0.15 M  $\text{Na}_3\text{Citrate}$ ), 0.1 percent sodium dodecylsulfate, for 30 min at 37°C prior to autoradiography. Positive plaques were re-screened with  $^{32}\text{P}$ -labeled probes derived from the second and third segments of the synthetic TNF- $\beta$  gene (13).

The TNF- $\alpha$  gene was isolated from a recombinant human fetal genomic library (27). Approximately  $7.5 \times 10^5$  plaques were screened by stringent hybridization (24) with the  $^{32}\text{P}$ -labeled TNF- $\alpha$  cDNA (14).

Phage DNA was prepared (27) and the hybridizing region was mapped with several restriction endonucleases, as outlined in figures 1 and 2.

DNA sequencing was performed by the dideoxy chain termination technique (28) after subcloning restriction fragments of the two TNF genes into the M13 cloning vectors mp8 and mp9 (29).

The 5' terminus of the TNF- $\beta$  transcript was determined by S1 nuclease mapping (30,31). Poly(A) RNA from mitogen-induced PBLs was annealed to a  $^{32}\text{P}$ -labeled, denatured StuI-BamHI fragment (286 bp, from position 663 to 949 of Figure 3), containing the beginning of the TNF- $\beta$  gene. The annealed mixture was treated with S1 nuclease to remove single stranded DNA and RNA. The size of the protected DNA fragment was measured on a sequencing gel along side a sequencing ladder of the same gene fragment.

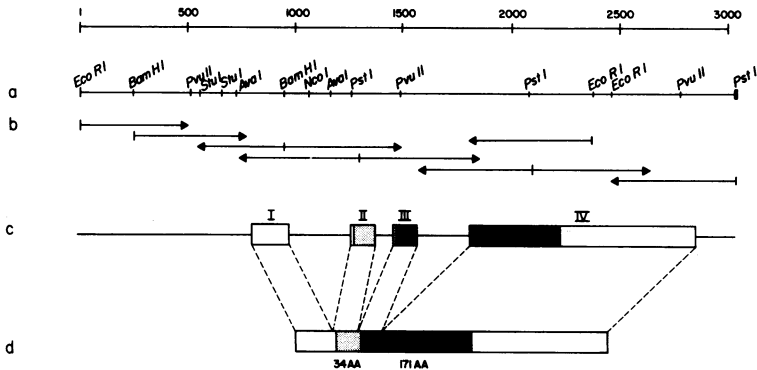
Chromosomal localization of the two human TNF genes was performed by Southern hybridization of DNA from human-murine somatic cell hybrids with the human TNF- $\alpha$  and TNF- $\beta$  cDNA probes. The human-murine somatic cell hybrids were previously constructed (32,33) and all were characterized for their human chromosome content by karyotyping (34) and by assaying assigned marker enzymes (35,36).

## RESULTS

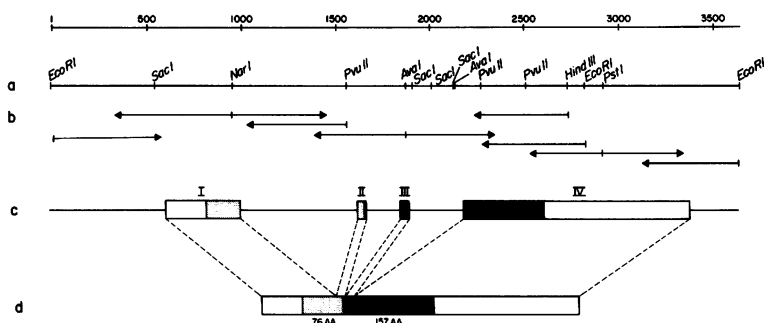
Isolation of the TNF- $\alpha$  and TNF- $\beta$  genes

The TNF- $\beta$  gene was isolated from a human genomic DNA- $\lambda$  library. Synthetic DNA was designed to encode the 34 amino-terminal residues of TNF- $\beta$  (see reference 13 for details). This DNA was subcloned, sequenced, and then used as a probe for hybridization under low stringency. Twenty-five hybridizing phage were identified and subsequently screened with two additional TNF- $\beta$  probes designed to code for residues 35-84 and 85-155 (the designed probe sequences were 70 percent homologous with the natural cDNA, 13). Five of the 25 recombinant phage hybridized with all three probes, and one of these,  $\lambda$ XB13, was characterized by restriction endonuclease analysis and Southern hybridization. All three TNF- $\beta$  probes hybridized with a 2.4 kbp EcoRI fragment. This fragment and an overlapping 950 bp PstI fragment (containing 3' untranslated sequences) were subcloned and sequenced as shown in Figure 1.

The human TNF- $\alpha$  gene was isolated from the human genomic DNA- $\lambda$  library described by Lawn *et al.* (27). A phage containing a 17.5 kb human genomic DNA fragment was identified using the entire TNF- $\alpha$  cDNA (14) as a probe.



**Figure 1.** The human TNF- $\beta$  gene structure and sequencing strategy. (a), Restriction endonuclease map of the TNF- $\beta$  gene derived from the recombinant bacteriophage  $\lambda$ XB13 DNA. (b), Arrows indicate the direction and extent of sequence analysis of each fragment. (c), The diagram shows a schematic representation of the TNF- $\beta$  gene as determined by DNA sequencing (see Figure 3). The mature TNF- $\beta$  coding region is represented by solid bars. The region encoding the signal sequence is marked by stippling and the open bars indicate the 5' and 3' untranslated regions. (d), Schematic representation of the TNF- $\beta$  mRNA. The putative signal sequence of 34 amino acids (34 AA) and the 171 amino acids (171 AA) of the mature coding region are noted as in (c). The bar graph at the top of the figure denotes the size of the TNF- $\beta$  gene in base pairs.



**Figure 2.** The human TNF- $\alpha$  gene structure and sequencing strategy. (a), Restriction endonuclease map of the TNF- $\alpha$  gene derived from the recombinant bacteriophage  $\lambda$ 5 DNA. (b), Arrows indicate the direction and extent of sequence analysis of each fragment. (c), The diagram shows a schematic representation of the TNF- $\alpha$  gene as determined by the DNA sequence of Figure 4. The mature TNF- $\alpha$  coding region is represented by solid bars. The region encoding the signal sequence is marked by stippling and open bars indicate the 5' and 3' untranslated regions. d, Schematic representation of the TNF- $\alpha$  mRNA. The putative signal sequence of 76 amino acids (76 AA) and the 157 amino acids of the mature coding region are noted. The size of the TNF- $\alpha$  gene is denoted in bp at the top of the figure.

The hybridizing region was subcloned as a 3.4 kb *Pst*I fragment and an overlapping 825 bp *Eco*RI fragment from phage  $\lambda$ 5, as shown in Figure 2. Southern blot analysis of total human genomic DNA suggests that TNF- $\alpha$  is encoded by a single gene, as is TNF- $\beta$  (13).

Nucleotide sequence analysis

The DNA sequences of the TNF- $\alpha$  and TNF- $\beta$  genes were determined by the dideoxy chain termination method as outlined in Figures 1 and 2. More than three kbp of sequence were obtained for both the TNF- $\beta$  gene (Figure 3) and the TNF- $\alpha$  gene (Figure 4). Each sequence was aligned with the previously determined, corresponding cDNA sequence (13,14).

An analysis of the TNF- $\beta$  gene sequence (Figure 3) reveals that the primary transcript contains three intervening sequences. The first intron (287 bp in length) interrupts the 5' untranslated region nine bp before the coding region. The second intron (86 bp in length) interrupts the signal sequence one residue before the mature sequence. The third intron (247 bp in length) interrupts codon 35 of mature TNF- $\beta$ . The sequences at the ends of each intron are homologous with consensus splice site sequences observed for many other genes (37). In particular, each intron begins with GT and ends with AG; a pyrimidine-rich region precedes the 3' end of each intron.



```

EcoRI
  1 GAATTCGGGGTATTCTACTCCGGGTGTCCAGGTTGTCTGCTACCCACCCAGCCTTCTGAGGCTCAAGCTGCACCAAGCCCCAGCTCCTCTCCCGCAGGACCCAACA
121 CAGGCTCAGGACTCAACACAGCTTTCCCTCCACCGCTTTCTCCTCCACGGACTCAGCTTTCTGAAAGCCCTCCACGTCTGTAGTCTCTATCTTTTCTCGCATCCTGCTGGAG
241 TTAGAAAGAAACAGACACAGACCTGGTCCCCAAAAGAAATGGAGCAATAGGTTTGAGGGGCATGGGGACGGGGTTCAGCTCCAGGGCTCTACACAAATCAGTCAGTGCCCAAG
361 AGACCCCCCTGGAAATCGGACAGGAGGATGGGGAGTGTAGGGGATCCTGTATGCTTGTTGTTCCCACTTTCCAATCCCGCCCGGCGATGGAGAAGAAACCGAGACAGAAG
481 TGCAGGGCCACTACCCTCCCTCCAGATGAGCTATGGGTTCTCCACCAGGAAGTTTCCGCTGGTGAATGATCTTTCCCGCCCTCCTCTCGCCCCAGGACATATATGGGCGAG
601 TTGTGGCACACCAGCCAGCAGCGCTCCCTCAGCAAGGACAGCAGGACCGAGCTAAGAGGGAGAGAAGCACTACAGACCCCTTGAACACCTCTCAGCCGCACATCCCTGTA

                                     -70
721 CAAGCTCAGCAGGTTCTCTCCCTCACATACTGACCCAGGGCTCACCCCTCTCCTCCCTGAAAGGACACC           met ser thr glu ser met ile arg asp val glu
                                        ATG AGC ACT GAA AGC ATG ATC CGG GAC GTG GAG
                                       -60           -50           -40
189 leu ala glu glu ala leu pro lys lys thr gly gly pro gln gly ser arg arg cys leu phe leu ser leu phe ser phe leu ile val
  CTG GCC GAG GAG GCG CTC CCC AAG AAG ACA GGG GGG CCC CAG GGC TCC AGG CGG TGC TTG TTC CTC AGC CTC TTC TCC TTC CTC ATC GTG

                                     -30           -20
919 ala gly ala thr thr leu phe cys leu leu his phe gly val ile gly pro gln arg glu glu
   GCA GGC GCC ACC ACG CTC TTC TGC CTG CTG CAC TTT GGA GTG ATC GGC CCC CAG AGG GAA GAG GTAGTGCTGCCAGCCTTCATCACTCTCCAC
1017 CCAAGGGGAAATGAGACCGCAAGAGGGAGAGATGGATGGGTAAGATGTGCGCTGATAGGAGGGATGAGAGAGAAAAACATGGAGAAGACGGGATGCAAGAAAGAGAT
1137 TGGCAAGAGATGGGAAGAGAGAGAGAAAGATGGAGAGACAGGATGTGGCACATGGAAGTGTCTACTAAGTGTGTATGGAGTGAATGAATGAATGAATGAATGAACAAGCAGATA
1257 TATAATAAGATATGGAGCAGATGTGGGTGTGAGAAGAGAGATGGGGGAAGAAACAAGTGAATGAATAAAGATGGAGACAGAAAGCGGAAATATGACAGCTAAGGAGAGAGA
1377 TGGGGAGATAGGAGAGAAAGATAGGTTCTGGCACAGAGACACTCAGGGAAGAGCTGTTGAATGCTGGAAGGTGAATACACAGATGAATGGAGAGAAAACCGAGACACT

                                                                           -10
1497 CAGGGCTAAGAGCGCAGCCAGCAGCGACGCTGTTCCCTTTAAGGGTGAATCCCTCGATGTTAACCATTCTCCTTCTCCCAACAG           phe pro arg asp leu ser leu
                                                                           TGC CCC AGG GAC CTC TCT CTA

                                      1
1610 ATC AGC CCT CTG GCC CAG GCA GTC A GTAAGTGTCTCAAACTCTTCTAATCTGGGTTTGGGTTTGGGGTAGGTTAGTACCGGTTAGGAAGCAGTGGGGAAAAT

                                          10
1720 TAAAGTTTGGCTCTGGGGAGGATGGATGGAGGTGAAAGTAGGGGGTATTTTCTAGGAAGTTTAAAGGCTCAGCTTTTCTTTCTCTCCTCTTCCAG           rg Ser Ser Ser
                                          GA TCA TCT TCT

                                          10
1833 Arg Thr Pro Ser Asp Lys Pro Val Ala His Val Val A
  CCG ACC CCG AGT GAC AAG CCT GTA GCC CAT GTT GTA G GTAAGAGCTCTGAGGATGTGTCTTGAAGACTGGAGGCTAGGATTTGGGGATTGAAAGCCGGCTGATGG
1939 TAGGCAGAATCTGGAGCAATGTGAGAAAGACTCGCTGAGCTCAAGGAAAGGGTGGAGGAACAGCACAGGCCCTTAGTGGGTACTCAGAAGCTATGCCAGGTGGGATGTGGATGACA

                                                                           1a Asn
2059 GACAGAGAGCAGGAAACCGSATGTGGGGTGGCGAGGCTCGAGGGCCAGGATGTGGAGAGTGAACCGACATGGCCACACTGACTCTCCTCCTCTCCTCCTCCCTCCAG           CA AAC

                                   20           30           40
2176 Pro Gln Ala Glu Gly Gln Leu Gln Trp Leu Asn Arg Arg Ala Asn Ala Leu Leu Ala Asn Gly Val Glu Leu Arg Asp Asn Gln Leu Val
  CCT CA A GCT GAG GGG CAG CTC CAG TGG CTG AAC CCG CGG GCC AAT GCC CTC CTG GCC AAT GGC GTG GAG CTG AGA GAT AAC CAG CTG GTG

                                   50           60           70
2266 Val Pro Ser Glu Gly Leu Tyr Leu Ile Tyr Ser Gln Val Leu Phe Lys Gly Gln Gly Cys Pro Ser Thr His Val Leu Leu Thr His Thr
  GTG CCA TCA GAG GGC CTG TAC CTC ATC TAC TCC CAG GTC CTC TTC AAG GGC CAA GGC TGC CCC TCC ACC CAT GTG CTC CTC ACC CAC ACC

                                   80           90           100
2356 Ile Ser Arg Ile Ala Val Ser Tyr Gln Thr Lys Val Asn Leu Leu Ser Ala Ile Lys Ser Pro Cys Gln Arg Glu Thr Pro Glu Gly Ala
  ATC AGC CAC ATC GCC CTC TCC TAC CAG ACC AAG GTC AAC CTC CTC TCT GCC ATC AAG AGC CCC TGC CAG GAG GAG ACC CCA GAG GGG GCT

                                   110           120           130
2446 Glu Ala Lys Pro Trp Tyr Glu Pro Ile Tyr Leu Gly Gly Val Phe Gln Leu Glu Lys Gly Asp Arg Leu Ser Ala Glu Ile Asn Arg Pro
  GAG GCC AAG CCC TGG TAT GAG CCC ATC TAT CTG GGA GGG GTC TTC CAG CTG GAG AAG GGT GAC CGA CTC AGC GCT GAT GAT ATC AAT CGG CCC

                                   140           150
2536 Asp Tyr Leu Asp Phe Ala Glu Ser Gly Gln Val Tyr Phe Gly Ile Ile Ala Leu
  GAC TAT CTC GAC TTT GCC GAG TCT GGG CAG GTC TAC TTT GGG ATC ATT GCC CTG TGA GGAGGCAGAACCTCCAACCTTCCCAAGCGCTCCCTCCGCCCCA
2636 ATCCCCTTATTACCCCTCCTTCAGACACCCTCAACCTCTCTGCTCAAAAAGAGAATGGGGCTTAGGGTCGAACCAAGCTTAGAACCTTTAAGCAACAAGACCACCACTCGAAA
2756 CCTGGGATTCAGGAATGTGGCTGACAGTGAAGTGTGGCAACCACCTAAGAACTCAAATGGGGCTCAGAACCTCAGAGGCTTACAGCTTGAATCCCTGACATCTGGAATCTG
2876 AGACAGGGAGCCCTTGGTTCTGGCAAGATGCTGACAGACTTGAAGAGACTCACCTAGAAAATGACACAAGTGGACCTTAGGGCTCTCCTCCAGATGTTCCAGACTTCTCTGAG
2996 ACACGAGGCCAGCCCTCCCACTGGAGCCAGCTCCCTCTATTTATGTGTGACTGTGATTTATTTATTTATTTATTTATTTATTTATTTACAGATGAATGATTTATTTGGGAGCCG
3116 GGGATCTCGGGGACCCCAAGTGGAGCTGCTTGGCTGAGACTGCTTTCCTGGTCAAGATGCTTTCTGGAACCGAGCTGAAACAAAGCTGTCCCTCATGTAGCCCTGTGGCTCTTTTTGAT
3236 ATGTTTTTAAATATTATCTGATTAAAGTGTCTAACCAACTGCTGATTTGGTGACCACTGTCACTATTGCTGAGCCTGTCTGCCCAAGGGAGTGTGCTGTAATGCCCTACTAT
3356 TCAGTGGCGAGAAATAGGTTGCTTGAAGAAAACATGGTCTCCTCTTGGAAATTAATTCTGCATCTGCCTCTTCTTGGTGGGAGAAAGCCCTTCAAGTCTCTCTCCACAGGCT
3476 TTAGATCCCTCGACCCAGCTCCATCTTAGACTCTCAGGCGCTTGAGACCTTACATAAACAAGAGCCCAACAGATATTTCCCCTCCTCCCGAACAACAGGACCTGAACCTAATATCC
3596 TCTCCCTCAGGGCATGGGAATTTCAACTCTGGGAATTC
EcoRI
  
```

Figure 4. The human TNF- $\alpha$  gene sequence. Positive numbers denote mature TNF- $\alpha$  coding region amino acids and negative numbers denote putative signal sequence amino acids. Stars indicate the putative site of initiation of transcription (residue 615) and poly(A) addition (residue 3382). The TATA box is overlined and AATAAA sequence is underlined.

The start of the TNF- $\beta$  mRNA (Figure 3) was identified by S1 nuclease mapping (30,31). Poly(A) RNA isolated from mitogen activated PBL was annealed to a  $^{32}\text{P}$ -labeled fragment containing the beginning of the TNF- $\beta$  gene. Four fragments were protected by the RNA (data not shown) which differed in length by a few bases, suggesting that the 5' end of the TNF- $\beta$  mRNA may be heterogeneous or that S1 nuclease may not precisely trim the single stranded DNA. These results indicate that transcription begins at approximately nucleotide 818. Consequently, the 5'-untranslated region of the TNF- $\beta$  mRNA is approximately 170 bases in length. The predicted length of the processed TNF- $\beta$  mRNA is thus 1420 bases without poly(A) addition. This agrees well with the observed size (14S) determined by Northern hybridization (13).

The TNF- $\alpha$  gene sequence (Figure 4) also contains three intervening sequences. Only one of these introns, however, is in a position homologous with the TNF- $\beta$  gene. The first intron (607 bp in length) in the TNF- $\alpha$  gene interrupts the precursor sequence after the codon for residue 62. The second intron (187 bp in length) interrupts the gene at the start of the mature TNF- $\alpha$  coding sequence, in the codon for the second residue. The third intron (301 bp) occurs at a homologous position with respect to the TNF- $\beta$  gene, in the codon for residue 18. The intron-exon junctions are similar to consensus splice site sequences observed for other eucaryotic genes (37). Recently, Shirai and coworkers (38) reported the sequence of the human TNF- $\alpha$  gene isolated from the same genomic DNA source described here. Their sequence is 325 bp shorter at the 5' end than reported here, but otherwise in agreement.

The 5'-untranslated region of TNF- $\alpha$  is approximately 180 nucleotides and most likely begins at nucleotide 615 (Figure 4). This was deduced from the limited homology observed in the putative 5' control regions of TNF- $\alpha$  and TNF- $\beta$  (see below). We have also isolated a TNF- $\alpha$  cDNA from the RPMI 1788 cell line which begins at nucleotide 616 (unpublished results). The predicted length of TNF- $\alpha$  mRNA is 1672 bases which compares favorably with the size (18S) observed by Northern analysis (14).

The most significant region of homology between the two TNF genes is found in the last exon, which codes for 80 percent of secreted TNF- $\beta$  and 89 percent of secreted TNF- $\alpha$ . Previous comparison of TNF- $\alpha$  and TNF- $\beta$  suggested that the proteins were 28 percent homologous (14). By rearranging the positions of gaps, this homology can be increased to 35 percent (17). This maximal homology, reflected at the nucleotide level, is 56 percent in the

|               |  |
|---------------|--|
| TNF- $\beta$  | TTTCCAGAACTCAGTCGCCTGAACCCCAAGCCTGTGGTT        |
| TNF- $\alpha$ | CTTGTGTGTCCTCAACTTCCAAATCCCGCCCGCCGCGAT        |
| CONSENSUS     | -TT---G--C-CA-----C--AA-CCCC--CC---G-T         |
| TNF- $\beta$  | CTCTCCTAGGCCTCAGCCTTCTGCTTTGACTGAAACA          |
| TNF- $\alpha$ | GGAGAAGAAACGAGACAGAAGGTGCAAGGCCCCTACTACCG      |
| CONSENSUS     | -----A--CC---C-----TGC-----C---A-C-            |
| TNF- $\beta$  | GCAGTATCTTCTAAGCCCTGGGGCTTCCCGGGCCCGAG         |
| TNF- $\alpha$ | <u>CTTCTC</u> CAGATGAGCTCATGGTTTCTCCACCAAGGAAG |
| CONSENSUS     | -----C---T-AGC-C--GGG--TC-CC-----AG            |
| TNF- $\beta$  | CCCCGACCTAGAACCCTGGCGCTGCCTGCCACGCTGCCAC       |
| TNF- $\alpha$ | TTTTCCGCTGTTGAATGATTCTTTCCCGCCCTCCTCTC         |
| CONSENSUS     | -----CT-G-----CT--C--C--C-----C-C              |
| TNF- $\beta$  | <u>TGCCGTTCTC</u> TATAAAGGGACCTGAGCGTCCGGGCCCA |
| TNF- $\alpha$ | GGCCAGGGACATATAAGGCAGTTGTTGGCAC--ACCCA         |
| CONSENSUS     | --CC-----TATAAGG-A-TG--G--C--CCCA              |

5' CAP

Figure 5. Comparison of the putative promoter regions of the TNF- $\alpha$  and TNF- $\beta$  genes. The sequences represent 200 bases upstream from the putative cap site, shown at the 3' end. The TNF- $\beta$  sequence begins with residue 619 of Figure 3 and the TNF- $\alpha$  sequence begins with residue 418 of Figure 4. The underlined nucleotides indicate the region of greatest homology between the two genes (15 out of 16 identical residues).

coding region of the fourth exon. The 3' untranslated sequences of TNF- $\alpha$  and - $\beta$  (793 and 633 bases, respectively) are 43 percent homologous, and each contain a large stretch of A-T residues (position 2597-2620 in the TNF- $\beta$  gene, and position 3032-3075 in the TNF- $\alpha$  gene).

Other regions of the TNF- $\alpha$  and - $\beta$  transcripts do not appear to be significantly homologous. The introns in particular are quite dissimilar (less than 35 percent homology); even the third introns, which occur in a homologous position in these genes, exhibit no obvious homology. The coding regions of the signal sequences are not homologous, an expected result considering the large size differences of the precursors (76 residues for TNF- $\alpha$  and 34 residues for TNF- $\beta$ ). The 5'-untranslated sequences also do not appear to be significantly homologous.

#### Putative control regions

The 5' putative control regions of the TNF- $\alpha$  and TNF- $\beta$  genes have several stretches of exact nucleotide homology and an overall homology of 35 percent for the sequences shown in Figure 5. On a completely random basis, 25 percent of the nucleotides should be coincident. Both of the TNF genes have the same Goldberg-Hogness TATAAA sequence (39,40). Presumably involved in promoting transcription initiation, this sequence is located 27 bp and 28 bp 5' of the putative cap site in the TNF- $\alpha$  and TNF- $\beta$  genes, respectively. Another consensus sequence (GG<sup>C</sup>CAATCT) thought to be



TABLE 1  
SEGREGATION OF TNF- $\alpha$  AND TNF- $\beta$  GENES WITH HUMAN CHROMOSOMES IN SOMATIC CELL HYBRIDS

| HYBRID        | HYBRIDIZATION WITH  |                    | HUMAN CHROMOSOMES |    |    |    |    |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | TRANSLOCATION |      |
|---------------|---------------------|--------------------|-------------------|----|----|----|----|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---------------|------|
|               | TNF- $\alpha$ PROBE | TNF- $\beta$ PROBE | 1                 | 2  | 3  | 4  | 5  | 6 | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |               | X    |
| TSL-2         | +                   | +                  | -                 | +  | -  | -  | +  | + | -  | -  | +  | -  | +  | -  | -  | -  | -  | -  | -  | +  | -  | +  | +  | -  | +             | 17/3 |
| ATR-13        | +                   | +                  | +                 | +  | +  | -  | +  | + | +  | -  | +  | -  | +  | +  | +  | +  | +  | +  | +  | -  | +  | -  | -  | -  | -             | 5/X  |
| NSL-9         | -                   | -                  | -                 | -  | -  | +  | -  | - | +  | -  | +  | -  | +  | +  | +  | +  | +  | +  | +  | -  | -  | +  | +  | +  | 17/9          |      |
| NSL-7         | +                   | +                  | -                 | -  | -  | -  | +  | - | -  | -  | -  | -  | +  | +  | +  | +  | +  | +  | +  | -  | +  | +  | +  | -  | 12p-          |      |
| NSL-15        | +                   | +                  | -                 | +  | -  | +  | +  | - | +  | +  | -  | -  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  |               |      |
| JSR-17S       | -                   | -                  | +                 | -  | +  | -  | -  | + | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | -  | +  | +  | +  | -  | 7q-           |      |
| EXR-5CSAZ     | +                   | +                  | +                 | -  | +  | +  | +  | + | +  | +  | -  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | X/11          |      |
| WIL-6         | +                   | +                  | -                 | +  | +  | +  | +  | + | +  | +  | -  | +  | -  | -  | +  | -  | -  | +  | -  | +  | +  | +  | +  | -  |               |      |
| WIL-8X        | -                   | -                  | -                 | -  | +  | +  | +  | - | +  | -  | -  | +  | +  | +  | -  | +  | -  | -  | +  | +  | +  | +  | +  | -  |               |      |
| WIL-7         | +                   | +                  | -                 | +  | +  | -  | +  | - | +  | -  | +  | +  | -  | +  | +  | -  | +  | -  | +  | +  | -  | +  | -  | +  |               |      |
| JWR-26C       | +                   | +                  | -                 | +  | +  | +  | +  | - | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | +  | -  | 1p-           |      |
| JWR-22H       | +                   | +                  | -                 | -  | -  | +  | -  | - | -  | -  | +  | -  | -  | -  | -  | -  | -  | -  | -  | +  | +  | +  | -  | -  | 2/1           |      |
| REW-11        | -                   | -                  | -                 | -  | -  | +  | -  | - | -  | -  | -  | +  | -  | +  | -  | -  | +  | -  | -  | -  | -  | +  | +  | -  |               |      |
| XER-11        | +                   | +                  | +                 | -  | +  | +  | -  | + | +  | +  | -  | +  | -  | +  | -  | -  | +  | +  | +  | +  | +  | +  | +  | +  | 11/X X/11     |      |
| XTR-22        | +                   | +                  | -                 | +  | +  | +  | +  | - | +  | -  | +  | -  | -  | -  | -  | -  | -  | -  | -  | +  | +  | +  | +  | -  | X/3           |      |
| % DISCORDANCE |                     |                    | 60                | 27 | 53 | 33 | 47 | 6 | 40 | 40 | 67 | 33 | 60 | 47 | 67 | 40 | 53 | 80 | 40 | 27 | 40 | 53 | 33 | 60 | 47            |      |

important in modulating RNA polymerase II transcription (41) and usually located 70 to 80 bp upstream of the cap site, is not readily identifiable in either gene. The longest stretch of exact homology (underlined, but not aligned in Figure 5) begins at residue 775 of the TNF- $\beta$  gene (just before the TATAAA sequence) and residue 489 of the TNF- $\alpha$  gene (100 bp before the TATAAA sequence); fifteen of sixteen residues are identical when these positions are compared.

#### Chromosomal localization

Southern blot analysis of DNA from 15 independently derived human-murine somatic cell hybrids was performed with both TNF cDNA probes. Hybridization with the human TNF- $\beta$  probe to an EcoRI digest of DNA from these hybrids revealed a human TNF- $\beta$  band at 2.4 kbp and a murine TNF- $\beta$  band at 3.4 kbp. As shown in Table 1, the TNF- $\beta$  probe hybridized with the 2.4 kbp human band only in somatic cell hybrids containing human chromosome 6.

Since the murine TNF- $\alpha$  gene hybridized at the same position (2.8 kbp) as the human TNF- $\alpha$  gene using EcoRI digested DNA, HindIII digestions were performed. Two bands were detected of greater than 20 kbp and 2.8 kbp for murine and human TNF- $\alpha$ , respectively, when hybridized with the human TNF- $\alpha$  cDNA probe. As observed for TNF- $\beta$ , the TNF- $\alpha$  probe hybridized with DNA from hybrids containing human chromosome 6. Both probes hybridized weakly to DNA from one hybrid which did not contain chromosome 6 by karyotype analysis; however, the overall discordance for chromosome 6 localization is much lower

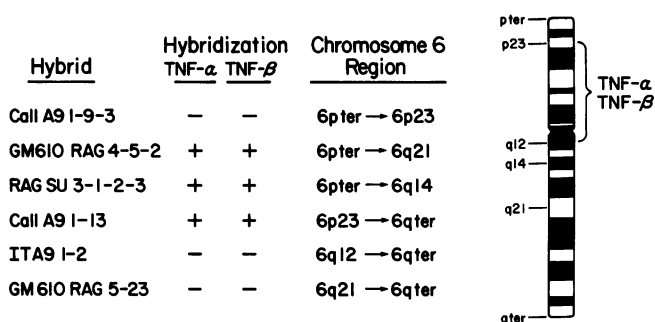


Figure 6. Regional localization of the TNF- $\alpha$  and TNF- $\beta$  genes to human chromosome 6. The chromosome 6 banding pattern is redrawn from reference 48.

than for all other chromosomes (6 percent compared with 27-80 percent). This suggests that both genes are indeed encoded by chromosome 6.

Analysis of human-murine hybrids which contain fragments of human chromosome 6 provide information on the regional localization of the TNF- $\alpha$  and - $\beta$  genes. As shown in Figure 6, both genes hybridize with cell hybrid DNA containing the 6p23 $\rightarrow$ 6q12 segment of chromosome 6. This encompasses approximately one-third of this chromosome. The two genes have thus been mapped to a region containing about 2 percent of the human genome, indicating that these genes are closely linked.

#### DISCUSSION

TNF- $\alpha$  and TNF- $\beta$  are cytotoxic proteins with similar biochemical characteristics and biological activities. Their gene structures share certain similarities but also contain distinctive differences. Both genes are of similar size, with a primary transcript of 2762 bp for TNF- $\alpha$  and 2038 bp for TNF- $\beta$ . Each gene contains three intervening sequences, but only the third intron is in a homologous position. Both genes share homologous regions in the fourth exon (56 percent in the coding region) and perhaps in the putative promoter region (35 percent). In contrast, these genes have distinctive, non-homologous sequences in the first three exons and all introns. The higher degree of nucleotide homology in the last exon is reflected in the overall homology of the secreted proteins, since the last exon codes for greater than 80 percent of mature TNF- $\alpha$  and - $\beta$ . Based on this extensive nucleotide homology, it appears likely that the last exon of the two genes was derived from a common ancestral sequence. Also, it is highly

probable that the conserved amino acids encoded by the last exon are necessary for their shared cytotoxic activities. This analysis supports the exon shuffling theory of Gilbert (42), which proposes that exons coding for important protein domains may be rearranged to mediate the rapid evolution of new protein sequences.

The differences in the signal sequences and the gene structures which encode them may be important for their observed mode of action. TNF- $\beta$  is secreted from T-lymphocytes one to two days after mitogenic stimulation (22); these kinetics are similar to other lymphokines derived from T-cells, such as IFN- $\gamma$  (43). The TNF- $\beta$  signal sequence is characteristic of most signal sequences in that it is short, very hydrophobic and has charged residues near the amino terminus (44,45). TNF- $\alpha$ , on the other hand, has a very long precursor sequence of 76 residues containing both hydrophilic and hydrophobic regions. Interleukin-1, like TNF- $\alpha$ , is a product of activated macrophages, and also contains a long precursor sequence of 114 amino acids (46). The precursor sequence of TNF- $\alpha$  may be important for the directed release of this cytolytic protein; this may provide the motile macrophage with a mechanism for secreting TNF- $\alpha$  at a specifically desired site.

The limited homology of the putative promoter regions of the TNF genes (Figure 5) may aid in explaining the induction of these genes. Both genes are induced by similar mitogenic stimuli, such as staphylococcal enterotoxin or phorbol esters (13,14), and conserved nucleotide regions may be essential for a common induction. However, the kinetics of induction are quite different: TNF- $\alpha$  transcription begins within two hours of mitogenic stimulation, whereas no TNF- $\beta$  mRNA is detectable until after eight hours (22). The roles that cell type and DNA sequence play in this regulation remain to be determined.

Both genes are encoded by chromosome 6. This close linkage and limited nucleotide and protein homology imply that these genes are ancestrally related. This is similar to IFN- $\beta$  and the IFN- $\alpha$  gene family, which share approximately 30 percent amino acid homology (47) and are closely linked on human chromosome 9 (48). Interestingly, the HLA genes of the major histocompatibility complex map to the same region of chromosome 6 as the TNF- $\alpha$  and TNF- $\beta$  genes (49); this adjacent genetic localization may be useful for a coordinate regulation of immune system gene products.

#### ACKNOWLEDGMENTS

The authors wish to thank Jeanne Arch for preparation of the manuscript

and Figures 3 and 4. Analysis of the chromosomal localization was initiated at Roswell Park Memorial Institute, Buffalo, New York, and we thank Dr. Thomas Shows for the human-murine somatic cell hybrids used in this study. We also thank Dr. Karl Heinz Grzeschik of the University of Munster for human-murine hybrids used in the regionalization to chromosome six. This research was supported by Genentech, Inc. and grants from the National Cancer Institute (to A.Y.S.) and the March of Dimes (to S.L.N.).

\*Present address: Molecular Therapeutics, Inc., 400 Morgan Lane, West Haven, Connecticut 06516.

### REFERENCES

1. Ruddle, N.H., Powell, M.B. and Conta, B.S. (1983) *Lymphokine Res.* 2, 23-31.
2. Granger, G.A., Yamamoto, R.S., Devlin, J.J. and Klostergaard, J. (1982). *Lymphokine Res.* 1, 45-49.
3. Ruff, M.R. and Gifford, G.E. (1981) in *Lymphokines Vol. 2*, Pick, E., ed., pp. 235-275, Academic, New York.
4. Carswell, E.A., Old, L.J., Kassel, R.L., Green, S., Fiore, N., and Williamson, B. (1975) *Proc. Natl. Acad. Sci.* 72, 3666-3670.
5. Evans, C.H. (1982) *Canc. Immunol. Immunother.* 12, 181-190.
6. Rundell, J.O. and Evans, C.H. (1981) *Immunopharmacology* 3, 9-18.
7. Ruff, M.R. and Gifford, G.E. (1981) *Infect. Immun.* 31, 380-385.
8. Williamson, B.D., Carswell, E.A., Rubin, B.Y., Prendergast, J.S., and Old, L.J. (1983) *Proc. Natl. Acad. Sci.* 80, 5397-5401.
9. Williams, T.W. and Bellanti, J.A. (1983) *J. Immunol.* 130, 518-520.
10. Stone-Wolff, D.S., Yip, Y.K., Kelker, H.C., Le, J., Henriksen-Destefano, D., Rubin, B.Y., Rinderknecht, E., Aggarwal, B.B., and Vilcek, J. (1984) *J. Exp. Med.* 159, 828-843.
11. Lee, S.H., Aggarwal, B.B., Rinderknecht, E., Assisi, F., and Chiu, H. (1984) *J. Immunol.* 133, 1083-1086.
12. Powell, M.B., Conta, B.S., Horowitz, M., and Ruddle, N.H. (1985) *Lymphokine Res.* 4, 13-26.
13. Gray, P., Aggarwal, B.B., Benton, C.V., Bringman, T.S., Henzel, W.J., Jarrett, J.A., Leung, D.W., Moffat, B., Ng, P., Svedersky, L.P., Palladino, M.A., and Nedwin, G.E. (1984) *Nature* 312, 721-724.
14. Pennica, D., Nedwin, G.E., Hayflick, J.S., Seeburg, P.H., Derynck, R., Palladino, M.A., Kohr, W.J., Aggarwal, B.B., and Goeddel, D.V. (1984) *Nature* 312, 724-729.
15. Aggarwal, B.B., Moffat, B. and Harkins, R.N. (1984) *J. Biol. Chem.* 259, 686-691.
16. Aggarwal, B.B., Henzel, W.J., Moffat, B., Kohr, W.J., and Harkins, R.N. (1985) *J. Biol. Chem.* 260, 2334-2344.
17. Aggarwal, B.B., Kohr, W.J., Hass, P.E., Moffat, B., Spencer, S.A., Henzel, W.J., Bringman, T.S., Nedwin, G.E., Goeddel, D.V., and Harkins, R.N. (1985) *J. Biol. Chem.* 260, 2345-2354.
18. Matthews, N. (1978) *Br. J. Cancer* 38, 310-315.
19. Currie, G.A. (1978) *Nature (London)* 273, 758-759.
20. Reed, W.P. and Lucas, Z.J. (1975) *J. Immunol.* 115, 395-404.
21. Svedersky, L.P., Nedwin, G.E., Goeddel, D.V., and Palladino, M.A. (1985) *J. Immunology*, in press.

22. Nedwin, G.E., Svedersky, L.P., Bringman, T.S., Palladino, M.A. and Goeddel, D.V. (1985) *J. Immunology*, in press.
23. Gitschier, J., Wood, W.I., Goralka, T.M., Wion, K.L., Chen, E.Y., Eaton, D., Vehar, G.A., Capon, D.J., and Lawn, R.M. (1984) *Nature* 312, 326-330.
24. Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180-182.
25. Gray, P.W. and Goeddel, D.V. (1983) *Proc. Natl. Acad. Sci. USA* 80, 5842-5846.
26. Ullrich, A., Berman, C.H., Dull, T.J., Gray, A., and Lee, J.M. (1984) *EMBO J.* 3, 361-364.
27. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) *Cell* 15, 1157-1174.
28. Smith, A.J.H. (1980) *Meth. Enzym.* 65, 560-580.
29. Messing, J. (1982) in *Genetic Engineering: Principles and Methods*, Hollaender, A. and Setlow, J., Eds., Plenum, New York, volume 4.
30. Berk, A.J. and Sharp, P.A. (1977) *Cell* 12, 721-732.
31. Weaver, R.F. and Weissmann, C. (1979) *Nucleic Acids Research* 7, 1175-1193.
32. Davidson, R.L. and Gerald, P.S. (1976) *Somat. Cell Genet* 2, 165.
33. Klebe, R.J., Chen, T.R., and Ruddle, F.H. (1970) *J. Cell Biol.* 45, 74.
34. Shows, T.B., Brown, J.A., Haley, L.L., Byers, M.G., Eddy, R.L., Cooper, E.S., and Goggin, A.P. (1978) *Cytogenet. Cell Genet.* 21, 99-104.
35. Shows, T.B., Sakaguchi, A.Y., and Naylor, S.L. (1982) *Adv. Hum. Genet.* 12, 341-452.
36. The Human Gene Mapping Workshop VI. (1982) *Cytogenet. Cell Genet.* 32, 1.
37. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) *Proc. Natl. Acad. Sci.* 75, 4853-4857.
38. Shirai, T., Yamaguchi, H., Ito, H., Todd, C.W., and Wallace, R.B. (1985) *Nature* 313, 803-806.
39. Goldberg, M. (1979) Thesis, Stanford University.
40. Gannon, F., O'Hare, Perrin, F., LePennec, J.P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B., and Chambon, P. (1979) *Nature* 278, 428-434.
41. Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) *Nucleic Acids Research* 8, 127-142.
42. Gilbert, W. (1978) *Nature* 271, 501.
43. Gray, P.W., Leung, D.W., Pennica, D., Yelverton, E., Najarian, R., Simonsen, C.C., Derynck, R., Sherwood, P.J., Wallace, D.M., Berger, S.L., Levinson, A.D., and Goeddel, D.V. (1982) *Nature* 295, 503-508.
44. Kreil, G.A. (1981) *Rev. Biochem.* 50, 317-348.
45. Watson, M.E.E. (1984) *Nucleic Acids Res.* 12, 5145-5164.
46. Lomedico, P.T., Gubler, U., Hellmann, C.P., Dukovich, M., Giri, J.G., Pan, Y-C.E., Collier, K., Semionow, R., Chua, A.O., and Mizel, S.B. (1984) *Nature* 312, 458-462.
47. Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Muramatsu, M., and Weissman, C. (1980) *Nature* 285, 547-549.
48. Owerbach, D., Rutter, W.J., Shows, T.B., Gray, P.W., Goeddel, D.V., and Lawn, R.M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3123-3127.
49. de Grouchy, J. and Turleau, C. (1984) *Clinical Atlas of Human Chromosomes*, 2nd ed., pp. 88-90, Wiley Medical, New York.