



(86) Date de dépôt PCT/PCT Filing Date: 1993/03/26  
 (87) Date publication PCT/PCT Publication Date: 1993/10/14  
 (45) Date de délivrance/Issue Date: 2005/03/01  
 (85) Entrée phase nationale/National Entry: 1994/09/29  
 (86) N° demande PCT/PCT Application No.: US 1993/002938  
 (87) N° publication PCT/PCT Publication No.: 1993/019777  
 (30) Priorité/Priority: 1992/03/30 (07/860,710) US

(51) Cl.Int.<sup>5</sup>/Int.Cl.<sup>5</sup> C12N 15/62, C12N 15/79, C12N 15/28,  
 C12N 15/25, A61K 38/19, A61K 38/17, C07K 14/715,  
 C07K 19/00  
 (72) Inventeur/Inventor:  
 SMITH, CRAIG A., US  
 (73) Propriétaire/Owner:  
 IMMUNEX CORPORATION, US  
 (74) Agent: SMART & BIGGAR

(54) Titre : PROTEINES DE FUSION CONTENANT UN RECEPTEUR DU FACTEUR DE NECROSE DES TUMEURS  
 (54) Title: FUSION PROTEINS COMPRISING TUMOR NECROSIS FACTOR RECEPTOR

(57) **Abrégé/Abstract:**

Fusion proteins comprise a tumor necrosis factor receptor (TNF-R) polypeptide and at least one additional polypeptide selected from an interleukin-1 receptor (IL-1R) and a second TNF-R polypeptide. One such fusion protein comprises one IL-1R and two TNF-R polypeptides. The fusion proteins have therapeutic use and may be produced via recombinant DNA technology.



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification<sup>5</sup> :</b> <b>A61K 37/66, C07H 21/04</b> <b>C07K 13/00, C12N 15/25, 15/28</b> <b>C12N 15/62, 15/79, 15/70</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 93/19777</b> <b>(43) International Publication Date:</b> 14 October 1993 (14.10.93) <div style="text-align: center; font-size: 1.5em; font-weight: bold;">2133326</div>
<b>(21) International Application Number:</b> PCT/US93/02938 <b>(22) International Filing Date:</b> 26 March 1993 (26.03.93)  <b>(30) Priority data:</b> 07/860,710                      30 March 1992 (30.03.92)                      US  <b>(71) Applicant:</b> IMMUNEX CORPORATION [US/US]; 51 University Street, Seattle, WA 98101 (US).  <b>(72) Inventor:</b> SMITH, Craig, A. ; 2405 5th West, Seattle, WA 98119 (US).  <b>(74) Agent:</b> SEESE, Kathryn, A.; Immunex Corporation, 51 University Street, Seattle, WA 98101 (US).		<b>(81) Designated States:</b> AU, CA, FI, JP, KR, NO, NZ, Euro- pean patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> FUSION PROTEINS COMPRISING TUMOR NECROSIS FACTOR RECEPTOR		
<b>(57) Abstract</b>  Fusion proteins comprise a tumor necrosis factor receptor (TNF-R) polypeptide and at least one additional polypeptide selected from an interleukin-1 receptor (IL-1R) and a second TNF-R polypeptide. One such fusion protein comprises one IL-1R and two TNF-R polypeptides. The fusion proteins have therapeutic use and may be produced <i>via</i> recombinant DNA technology.		

**TITLE**

Fusion Proteins Comprising Tumor Necrosis Factor Receptor

5

**BACKGROUND OF THE INVENTION**

A number of cytokines are known to bind to specific receptor proteins on the surface of target cells. Among the specific receptor proteins that have been identified are tumor necrosis factor receptors and interleukin-1 receptors. Much effort is being directed toward isolation and characterization of a number of receptors in order to study their physiological roles and to explore possible therapeutic uses. The binding of a particular target molecule by a soluble receptor administered to a patient may alleviate disorders mediated by the target molecule.

15 Tumor necrosis factor- $\alpha$  (TNF $\alpha$ , also known as cachectin) and tumor necrosis factor- $\beta$  (TNF $\beta$ , also known as lymphotoxin) are homologous mammalian endogenous secretory proteins capable of inducing a wide variety of effects on a large number of cell types. The great similarities in the structural and functional characteristics of these two cytokines have resulted in their collective description as "TNF." Complementary cDNA clones encoding TNF $\alpha$  (Pennica et al., *Nature* 312:724, 1984) and TNF $\beta$  (Gray et al., *Nature* 312:721, 1984) have been isolated, permitting further structural and biological characterization of TNF.

25 TNF proteins initiate their biological effect on cells by binding to specific TNF receptor (TNF-R) proteins expressed on the plasma membrane of a TNF-responsive cell. TNF $\alpha$  and TNF $\beta$  were first shown to bind to a common receptor on the human cervical carcinoma cell line ME-180 (Aggarwal et al., *Nature* 318:665, 1985). Hohmann et al. (*J. Biol. Chem.* 264:14927, 1989) reported that at least two different cell surface receptors for TNF exist on different cell types, although the relationship between these TNF-Rs is unclear. These receptors have an apparent molecular mass of about 75 - 80 kDa and about 55-60 kDa, respectively. In addition to cell surface receptors for TNF, soluble proteins from human urine capable of binding TNF have also been identified (Peetre et al., *Eur. J. Haematol.* 41:414, 1988; Seckinger et al., *J. Exp. Med.* 167:1511, 1988; Seckinger et al., *J. Biol. Chem.* 264:11966, 1989; UK Patent Application, Publ. No. 2 218 101 A to Seckinger et al.; Engelmann et al., *J. Biol. Chem.* 264:11974, 1989).

35

64166-33

2

Interleukin-1 $\alpha$  (IL-1 $\alpha$ ) and interleukin-1 $\beta$  (IL-1 $\beta$ ) are distantly related polypeptide hormones that play a central role in the regulation of immune and inflammatory responses. These two proteins act on a variety of cell types and have multiple biological activities. The biological activities ascribed to IL-1 $\alpha$  and IL-1 $\beta$  are mediated via at least two classes of plasma membrane bound receptors which bind both IL-1 $\alpha$  and IL-1 $\beta$ . The IL-1 receptors expressed on B cells (referred to herein as type II IL-1 receptors) are different from IL-1 receptors detected on T cells and other cell types (referred to herein as type I IL-1 receptors).

#### SUMMARY OF THE INVENTION

The present invention is directed to receptors that are fusion proteins in which:

- (1) a first tumor necrosis factor receptor (TNF-R) polypeptide is covalently linked to a second TNF-R polypeptide, or
- (2) at least one TNF-R polypeptide is covalently linked to at least one interleukin-1 receptor (IL-1R) polypeptide.

The receptors preferably are produced as fusion proteins via recombinant DNA technology. The present invention provides fusion proteins comprising, as one of at least two biologically active polypeptide components, a TNF-R polypeptide. One fusion protein of the present invention comprises two TNF-R polypeptides, preferably joined via a peptide linker.

In another embodiment of the invention, the fusion protein comprises TNF-R and IL-1R. The fusion protein

64166-33

3

preferably comprises two TNF-R polypeptides and either one or two IL-1R polypeptides.

The present invention also provides isolated DNA sequences encoding the fusion proteins, recombinant  
5 expression vectors comprising such DNA sequences, host cells containing the expression vectors, and processes for producing the recombinant fusion proteins by culturing the host cells. Pharmaceutical compositions comprising a purified fusion protein as described above and a suitable  
10 diluent, carrier, or excipient are also provided by the present invention. Such compositions are useful in therapy, diagnosis, and assays for conditions mediated by tumor necrosis factor or interleukin-1.

#### BRIEF DESCRIPTION OF THE DRAWINGS

15 Figure 1 presents a restriction map for a type I human IL-1R cDNA clone. The sites at which certain restriction enzymes cleave the cDNA are shown.

Figure 2 depicts the partial cDNA sequence and derived amino acid sequence of a human TNF-R clone.  
20 Nucleotides are numbered from the beginning of the 5' untranslated region. Amino acids are numbered from the beginning of the signal peptide sequence. The apparent signal peptide sequence is represented by the amino acids -22 to -1. The N-terminal leucine of the mature TNF-R  
25 protein is underlined at position 1. The apparent transmembrane region from amino acids 236 to 265 is also underlined. The C-termini of various soluble TNF-Rs are marked with an arrow (↓). Cleavage sites for certain restriction endonucleases employed in constructing  
30 expression vectors are indicated.

64166-33

3a

Figure 3 depicts a plasmid vector comprising a DNA fragment encoding a fusion protein of the formula TNF-R-linker-TNF-R, constructed as described in example 11.

Figure 4 depicts a plasmid vector comprising a DNA  
5 fragment encoding a fusion protein of the formula IL-1R-linker-TNF-R-linker-TNF-R, constructed as described in example 12.

Figure 5 depicts three plasmid vectors that are intermediates in the construction of certain vectors of the  
10 present invention, as described in example 12.

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to receptors that are fusion proteins in which:

(1) a first TNF-R polypeptide is covalently  
15 linked to a second TNF-R polypeptide, or

(2) at least one TNF-R polypeptide is covalently linked to at least one IL-1R polypeptide.

The TNF-R and (when present) IL-1R polypeptide components may be attached to one another using any suitable  
20 technique for attaching one polypeptide to another. Cross-linking reagents and peptide linkers are among the linkage methods that may be employed. The TNF-R and IL-1R polypeptides are derived from mammalian species, preferably human.

25 The receptors preferably are produced as fusion proteins via recombinant DNA technology. The present invention provides fusion proteins comprising, as one component, a mammalian tumor necrosis factor receptor (TNF-R). One fusion protein of the present invention

64166-33

3b

comprises two TNF-R polypeptides and may be represented by the following formula:

TNF-R - linker - TNF-R

wherein the linker is a peptide linker.

In another embodiment of the invention, the fusion protein comprises TNF-R and a mammalian interleukin-1 receptor (IL-1R). In this embodiment of the invention, the fusion protein may comprise one TNF-R polypeptide and one IL-1R polypeptide.

5 Preferably, two TNF-R polypeptides and one IL-1R polypeptide are joined to form the fusion protein. The two TNF-R polypeptides preferably are adjacent to one another (as opposed to IL-1R being positioned between the two TNF-Rs) to enhance binding of TNF. Examples of such fusion proteins are represented by the following formulas:

10 TNF-R - linker - IL-1R

IL-1R - linker - TNF-R

TNF-R - linker - TNF-R - linker - IL-1R and

15

IL-1R - linker - TNF-R - linker - TNF-R

wherein each linker is a peptide linker. The N-terminus of each fusion protein is on the left side of each formula.

20 Each TNF-R polypeptide component of the fusion proteins is independently capable of binding tumor necrosis factor (TNF). Likewise, each IL-1R polypeptide employed in the fusion proteins is independently capable of binding interleukin-1 (IL-1). Including two adjacent TNF-R polypeptides in the fusion protein is advantageous in that the TNF binding affinity is increased compared to the binding of TNF by a  
25 single TNF-R polypeptide.

Peptide linkers that may be employed in the present invention separate TNF-R polypeptides (and IL-1R polypeptides, when present) from one another by a distance sufficient to ensure that each polypeptide properly folds into the secondary and tertiary structures necessary for the desired biological activity. The linker also should allow the  
30 extracellular domains of the TNF-R and IL-1R polypeptides to assume the proper spatial orientation to form a binding site for TNF or IL-1. The peptide linkers function as spacers, as opposed to the pharmaceutically active TNF-R and IL-1R polypeptide components of the fusion proteins. Suitable polypeptide linkers preferably (1) will adopt a flexible extended conformation, (2) will not exhibit a propensity for developing  
35 an ordered secondary structure which could interact with the functional protein domains, and (3) will have minimal hydrophobic or charged character which could



promote interaction with the functional protein domains. Typical surface amino acids in flexible protein regions include glycine (Gly), asparagine (Asn) and serine (Ser). Virtually any permutation of amino acid sequences containing Gly, Asn and Ser would be expected to satisfy the above criteria for a peptide linker sequence. Other near neutral amino acids, such as threonine (Thr) and alanine (Ala), may also be used in the linker sequence. Suitable peptide linkers generally comprise a chain of amino acids, preferably from 5 to 100 amino acids in length and most preferably from 10 to 20 amino acids in length. Examples of such linkers include, but are not limited to, (Gly<sub>4</sub>Ser)<sub>n</sub>, wherein n is 1-12, Gly<sub>4</sub>SerGly<sub>5</sub>Ser, and (Gly<sub>4</sub>SerGly<sub>5</sub>Ser)<sub>2</sub>.

### TNF-R and IL-1R Polypeptides

As used herein, the terms "interleukin-1 receptor" and "IL-1R" refer to proteins which are capable of binding interleukin-1 (IL-1) molecules and, in their native configuration as human plasma membrane proteins, play a role in transducing the signal provided by IL-1 to a cell. As used herein, the terms "TNF receptor" and "TNF-R" refer to proteins that are biologically active in that they are capable of binding tumor necrosis factor (TNF). Native membrane bound forms of the proteins also transduce a biological signal initiated by a TNF molecule binding to a cell. For fusion proteins comprising more than one TNF-R polypeptide, the TNF-R polypeptides may be identical or different. Likewise for IL-1R polypeptides.

Intact receptors generally include an extracellular domain which binds to a ligand, a hydrophobic transmembrane domain which remains embedded within the plasma membrane lipid bilayer, and a cytoplasmic or intracellular domain which is believed to deliver a biological signal to effector cells via a cascade of chemical reactions within the cytoplasm of the cell. The hydrophobic transmembrane domain and a highly charged region of the cytoplasmic domain immediately downstream of the transmembrane domain cooperatively function to halt transport of the IL-1 and TNF receptors across the plasma membrane. The extracellular domain of the TNF-R and IL-1R proteins disclosed herein is the N-terminal portion of the protein, from amino acid 1 to the amino acid immediately preceding the transmembrane region. The cytoplasmic domain is that portion of the protein that is located downstream of the transmembrane region.

Among the TNF-R polypeptides that may be employed as components of the inventive fusion proteins is a polypeptide comprising amino acids 1 to 439 of the sequence presented in Figure 2, which is a full length native TNF-R sequence. This TNF-R DNA and amino acid sequence is also presented in SEQ ID NOS: 1 and 2. When the signal sequence is desired, the TNF-R polypeptide may comprise amino

acids -22 to 439 of the figures 2A-2B sequence. The desirability of including the signal sequence depends on such factors as the position of the TNF-R polypeptide in the fusion protein and whether the intended host cells will process a mammalian signal sequence, as discussed below. The mature full-length native glycosylated form of this human TNF-R is a glycoprotein having a molecular weight of about 80 kilodaltons (kDa). As used throughout the specification, the term "mature" means a protein lacking a leader or signal sequence as may be present in full-length transcripts of a native gene. A protein may comprise a signal sequence when initially expressed. Cleavage of the signal sequence upon secretion of the protein from the cell yields the mature form of the protein.

Other suitable TNF-R polypeptides are described in European patent application publication number 422,339 (EP 422,339 hereinafter). Two suitable TNF-R polypeptides comprise arginine (Arg) as residue 174 but are otherwise identical to the above-described polypeptides comprising amino acids 1 to 439 or -22 to 439, respectively, of figures 2A-2B of the present application. A TNF-R amino acid sequence identical to that of figures 2A-2B (of the present application) except for substitution of arginine (Arg) for methionine (Met) at position 174 of the mature sequence is disclosed in EP 422,339 (see figure 39 therein).

The cDNA and encoded amino acid sequences of another useful TNF-R polypeptide are disclosed in figure 21 of EP 422,339. The coding region of the EP 422,339 figure 21 cDNA sequence, and the amino acid sequence encoded thereby, are presented as SEQ ID NOS:3 and 4 of the present application. Although referred to as a 30 kilodalton protein in EP 422,339, other molecular weights have been reported for this protein. A molecular weight of about 55 kilodaltons was reported by Loetscher et al. (*Cell* 61:351, 1990) and in EP 417,563, for example. Useful TNF-R polypeptides include those comprising amino acids 1 to 415 of the SEQ ID NO:4 sequence, or, when the signal sequence is desired, amino acids -40 to 415 of the SEQ ID NO:4 sequence. Methods for producing this TNF-R protein, either by purification from urine or from the medium of a culture of U937 cells, or by recombinant DNA technology, are described in EP 422,339. This TNF-R is characterized by an N-terminal sequence (for the mature form of the protein) of Asp-Ser-Val-Cys-Pro-Gln-, whereas the N-terminal sequence of the mature form of the TNF-R protein of figures 2A-2C is Leu-Pro-Ala-Gln-Val-Ala-.

In certain embodiments of the present invention, the TNF-R polypeptide is a soluble TNF-R polypeptide. Soluble TNF-R polypeptides lack at least part (preferably all) of the transmembrane region that promotes retention of the protein on the cell

surface. The soluble polypeptides generally also lack the charged region of the cytoplasmic domain (located immediately downstream of the transmembrane region) that contributes to retention on the cell surface. Preferably, the entire transmembrane region and cytoplasmic domain are deleted from the protein or substituted with hydrophilic amino acids to form soluble TNF-R. Soluble TNF-R is secreted from the cell and retains the desired biological activity.

Examples of soluble TNF-R polypeptides are those comprising amino acids 1 - x of the Figure 2A sequence, wherein x is the C-terminal amino acid and is selected from the group consisting of any one of amino acids 163 - 235 of Figure 2. Specific examples include polypeptides comprising the acids 1 - 163, 1 - 185, or 1 - 235 of Figure 2. The soluble TNF-R polypeptide may additionally comprise a signal sequence, e.g., amino acids -22 to -1 of Figure 2.

Additional examples of soluble TNF-R polypeptides are those comprising amino acids 1-184 or 1-182, -22-184, or -22-182 of the Figure 2 sequence. Such proteins may contain either methionine or arginine at position 174. Procedures for preparing examples of such TNF-R polypeptides include those described in examples 17 and 22 of EP 422,339.

The TNF-R protein shown in SEQ ID NO:4 comprises a signal peptide (designated amino acids -40 to -1) and a transmembrane region beginning with the valine residue at position 172. Preferred soluble forms of this TNF-R protein include those comprising amino acids -40-w or 1-w of SEQ ID NO:4, wherein w is an integer from 161-171 (i.e., any of amino acids 161 to 171 of SEQ ID NO:4 is the C-terminus). The use of oligonucleotide-directed *in vitro* mutagenesis to construct an expression vector encoding a biologically active TNF-R protein having amino acid 161 (asparagine) as the C-terminal amino acid is illustrated in example 7 of EP 422,339. Further, procedures for purifying naturally occurring soluble forms of both of the above-described TNF-R proteins (i.e., soluble forms of the SEQ ID NO:2 and the SEQ ID NO:4 proteins) from human urine have been described by Engelmann et al. (*J. Biol. Chem.* 265:1531, 1990).

Interleukin-1 receptors that may be employed as components of the inventive fusion proteins include polypeptides designated herein as type I IL-1R and type II IL-1R. Type I IL-1 receptors have been detected on T-cells and certain other cell types, while expression of type II IL-1 receptors on B cells has been reported. In the absence of any specific designation, the term "IL-1 receptor" as used herein refers collectively to type I and type II IL-1 receptors.

64166-33

8

Among the IL-1R polypeptides that may be employed in the present invention are the type I IL-1R polypeptides described in U.S. Patent Nos. 5,319,071 and 5,180,812; and in European Patent Publication No. 318,296. The DNA  
5 sequence of a cloned cDNA encoding a human type I IL-1R protein and the amino acid sequence encoded thereby are presented herein in SEQ ID NOS: 5 and 6. The protein contains 569 amino acids, of which 20 are an N-terminal signal peptide. The aspartic acid (Asp) residue at  
10 position 1 is the first amino acid of the mature protein. The transmembrane region includes amino acids 317 (His) through 336 (Tyr). The sequence of human IL-1R is also disclosed in Sims et al., *Proc Nat'l. Acad. Sci. USA* 86:8946 (1989).

15 As with the TNF-R polypeptides, soluble IL-1R polypeptides may be employed in the fusion proteins of the present invention. Soluble IL-1R polypeptides generally lack the transmembrane region and preferably lack the cytoplasmic domain as well. Soluble IL-1R proteins may also  
20 include part of the transmembrane region or the cytoplasmic domain, provided that the soluble IL-1R protein is capable of being secreted from the cell.

Examples of soluble type I IL-1R polypeptides include, but are not limited to, those comprising the amino  
25 acid sequence depicted as amino acids y-x of SEQ ID NO: 6, wherein x is 312-316 and y is -3 to 3. In other words, the N-terminal amino acid is selected from the Leu, Glu, Ala, Asp, Lys, and Cys residues at positions -3, -2, -1, 1, 2, and 3, respectively. The C-terminal amino acid of the  
30 soluble protein is selected from the Thr, Asn, Phe, Gln, and Lys residues at positions 312, 313, 314, 315, and 316, respectively. Preferred soluble IL-1R polypeptides include

64166-33

8a

those comprising amino acids 1-x of SEQ ID NO: 6, wherein x is 312-316.

Any of the above-described soluble type I IL-1R polypeptides may additionally comprise a signal sequence, e.g., the signal sequence shown as amino acids -20 to -1 in SEQ ID NO: 6. Alternatively, a different signal sequence functional in an intended host cell (e.g., a yeast signal sequence) may be employed, as discussed below.

Among the type II IL-1R polypeptides that may be employed in the present invention are those described in U.S. Patent No. 5,350,683; and in European Patent Publication No. 460,846. Native glycosylated human type II IL-1R proteins recovered from cell lysates generally have an apparent molecular weight of about 60-68 kilodaltons by SDS-PAGE. The DNA sequence of a cloned cDNA encoding a human type II IL-1R protein and the amino acid sequence

encoded thereby are presented herein in SEQ ID NOS:7 and 8. The protein comprises a 13-amino acid signal peptide. The transmembrane region includes amino acids 331 (Ala) through 356 (Met).

5 Soluble type II IL-1R proteins are derived by deleting a C-terminal portion of the protein that is not necessary for IL-1 binding, so that the protein is secreted from the cell. The cysteine residue at position 313 is believed to be necessary to maintain the tertiary structure of the type II IL-1R molecule and permit binding of IL-1. Examples of soluble type II IL-1R polypeptides thus include those in which the C-terminal amino acid is selected from any of amino acids 314-333. In other words, the soluble IL-1R  
10 may contain the amino acid sequence shown as amino acids 1-x of SEQ ID NO:8 wherein x is an integer from 314-333. Preferred examples of suitable soluble type II IL-1R polypeptides include, but are not limited to, those comprising amino acids 1-x of SEQ ID NO:8, wherein x is 330-333. In other words, the C-terminal amino acid of the soluble protein is selected from the Glu, Ala, Ser, and Ser residues at positions 330,  
15 331, 332, and 333, respectively. The soluble type II IL-1R polypeptides may additionally comprise a signal sequence, e.g., the signal sequence shown as amino acids -13 to -1 in SEQ ID NO:8. Alternatively, a different signal sequence functional in an intended host cell (e.g., a yeast signal sequence) may be employed, as discussed below.

20 Assay procedures described herein may be employed to confirm biological activity for additional soluble TNF-R and IL-1R polypeptides, beyond the particular examples set forth above. Soluble TNF-R and soluble IL-1R may be identified (and distinguished from their non-soluble membrane-bound counterparts) by separating intact cells which express the desired protein from the culture medium, e.g., by  
25 centrifugation, and assaying the medium (supernatant) for the presence of the desired protein. The culture medium may be assayed using procedures which are similar or identical to those described in the examples below. The presence of TNF-R and IL-1R in the medium indicates that the protein was secreted from the cells and thus is a soluble form of the desired protein.

30 The N- or C-terminus of the TNF-R or IL-1R polypeptides may vary according to such factors as the type of host cells employed when producing the fusion protein *via* recombinant DNA technology and the particular cells from which the protein is purified when non-recombinant TNF-R or IL-1R is employed. Such variations may be attributable to differential post-translational processing of the protein in various types of  
35 cells, for example. Variations in the N- or C-terminal sequence also may result from

2133326

the oligonucleotides chosen to reconstruct either terminus of the TNF-R or IL-1R encoding DNA sequence when constructing expression vectors.

Differential processing may result in mature TNF-R or IL-1R proteins having an N-terminal amino acid other than those shown at position 1 of SEQ ID NOS:2, 4, 6 and 8. For example, in certain host cells, post-translational processing will remove the methionine residue encoded by an initiation codon, whereas the methionine residue will remain at the N-terminus of proteins produced in other host cells. Further, the N- and C-termini have been known to vary for the same protein, depending on the source of the protein. In some cases, the deletion of amino acids at either terminus of the protein may be due to proteolysis, occurring either intracellularly or during purification. Varying N-termini may also result from cleavage of the signal peptide in certain host cells at a point other than between amino acids -1 and 1 of the disclosed sequences.

As described in examples 10 and 11 and figure 31 of EP 422,339, a non-recombinant mature TNF-R protein purified from human urine lacked two N-terminal amino acids found in a non-recombinant TNF-R protein purified from the human monocyte-like cell line U937. The N-terminal amino acid of the urine-derived TNF-R was the alanine residue at position 3 of SEQ ID NO:1. Engelmann et al. (*J. Biol. Chem.* 265:1531, 1990) disclose a protein that binds TNF and was purified from human urine in forms truncated to varying degrees at the N-terminus (see the abstract and page 1533). The N-terminal amino acid sequence of one protein species was Val-Ala-Phe-Thr-Pro-, which corresponds to amino acids 5-9 of SEQ ID NO:1. Other forms of the protein had either phenylalanine (amino acid 7 of SEQ ID NO:1) or threonine (amino acid 8 of SEQ ID NO:1) as the N-terminal amino acid.

The N- and C- termini of the TNF-R and IL-1R proteins may vary for reasons that include those discussed above. The N-terminal amino acid may, for example, be any of the amino acids at positions 1 to 5 of SEQ ID NOS: 2 or 4 for TNF-R or of SEQ ID NOS:6 or 8 for IL-1R. The C-terminus may be truncated deliberately during expression vector construction (e.g., in constructing vectors encoding soluble proteins as described above) or as a result of differential processing which may remove up to about five C-terminal amino acids, for example.

Additional TNF-R and IL-1R polypeptides that may be employed retain the desired biological activity but vary from the native sequences in that amino acid(s) are added to, deleted from, or substituted in the native sequence. The biological activity of such proteins can be confirmed using the assays described herein.

Derivatives of TNF-R and IL-1R that are within the scope of the invention also include various structural forms of the primary protein which retain biological activity.

Due to the presence of ionizable amino and carboxyl groups, for example, a TNF-R or IL-1R protein may be in the form of acidic or basic salts, or may be in neutral form. Individual amino acid residues may also be modified by oxidation or reduction.

5 The primary amino acid structure may be modified by forming covalent or aggregative conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives are prepared by linking particular functional groups to amino acid side chains or at the N- or C-termini. Naturally occurring variants may result from alternative RNA splicing events.

10 Other proteins that may be employed in the inventive fusion proteins include conjugates of TNF-R or IL-1R with other polypeptides, which may be produced by synthesis in recombinant culture as N-terminal or C-terminal fusions. For example, the conjugated peptide may be a signal (or leader) peptide sequence at the N-terminal region of the protein which co-translationally or post-translationally directs transfer of the protein from its site of synthesis to its site of function inside or outside of the cell  
15 membrane or wall (e.g., the yeast  $\alpha$ -factor leader). The fusion proteins can comprise peptides added to facilitate purification or identification of the fusion protein. Such peptides include, for example, poly-His or the antigenic identification peptides described in U.S. Patent No. 5,011,912 and in Hopp et al., *Bio/Technology* 6:1204,1988. One such peptide is the FLAG<sup>®</sup> peptide, Asp-Tyr-Lys-Asp-Asp-Asp-  
20 Asp-Lys (DYKDDDDK), which is highly antigenic and provides an epitope reversibly bound by a specific monoclonal antibody enabling rapid assay and facile purification of expressed recombinant protein. This sequence is also specifically cleaved by bovine mucosal enterokinase at the residue immediately following the Asp-Lys pairing. Fusion proteins capped with this peptide may also be resistant to intracellular  
25 degradation in *E. coli*. A murine hybridoma designated 4E11 produces a monoclonal antibody that binds the peptide DYKDDDDK in the presence of certain divalent metal cations (as described in U.S. Patent 5,011,912) and has been deposited with the American Type Culture Collection under accession no HB 9259.

30 The inventive fusion proteins comprise TNF-R or IL-1R with or without associated native-pattern glycosylation. TNF-R and IL-1R expressed in yeast or mammalian expression systems, e.g., COS-7 cells, may be similar or slightly different in molecular weight and glycosylation pattern than the native molecules, depending upon the expression system. Expression of recombinant proteins in bacteria such as *E. coli* provides non-glycosylated proteins. Proteins having inactivated N-glycosylation  
35 sites can be produced by oligonucleotide synthesis and ligation or by site-specific mutagenesis techniques. These mutant proteins can be produced in a homogeneous,



reduced-carbohydrate form in good yield using yeast expression systems. N-glycosylation sites in eukaryotic proteins are characterized by the amino acid triplet Asn-A<sub>1</sub>-Z, where A<sub>1</sub> is any amino acid except Pro, and Z is Ser or Thr. In this sequence, asparagine provides a side chain amino group for covalent attachment of carbohydrate. Such a site can be eliminated by substituting another amino acid for Asn or for residue Z, deleting Asn or Z, or inserting a non-Z amino acid between A<sub>1</sub> and Z, or an amino acid other than Asn between Asn and A<sub>1</sub>. Known procedures for inactivating N-glycosylation sites in proteins include those described in U.S. Patent 5,071,972 and EP 276,846. Examples of N-glycosylation sites in human type II IL-1R are amino acids 53-55, 59-61, 99-101, 206-208, and 264-266 in SEQ ID NO:8. Potential N-glycosylation sites are found in the type I IL-1R protein (SEQ ID NO:6) at amino acids 80-82, 173-175, 213-215, 229-231, 243-245, and 277-279. N-glycosylation sites are found at amino acid 171-173 and 358-360 of TNF-R in Figure 2.

Cysteine residues that are not essential for biological activity can be deleted or replaced with other amino acids to prevent formation of unnecessary or incorrect intramolecular disulfide bridges upon renaturation. The cysteine residue at position 178 in the Figure 2 TNF-R sequence can be deleted, for example. U.S. Patent 4,518,584 describes the use of site directed mutagenesis to delete or replace cysteine residues within a protein. Other approaches to mutagenesis involve modification of adjacent dibasic amino acid residues to enhance expression in yeast systems in which KEX2 protease activity is present. EP 212,914 discloses the use of site-specific mutagenesis to inactivate KEX2 protease processing sites in a protein. In order to preserve the biological activity of TNF-R and IL-1R, substitutions will preferably result in homologous or conservatively substituted sequences, meaning that a given amino acid residue is replaced by a residue having similar physiochemical characteristics. Examples of conservative substitutions include substitution of one aliphatic residue for another, such as Ile, Val, Leu, or Ala for one another, or substitutions of one polar residue for another, such as between Lys and Arg; Glu and Asp; or Gln and Asn. Other such conservative substitutions, for example, substitutions of entire regions having similar hydrophobicity characteristics, are well known. Moreover, particular amino acid differences between human, murine and other mammalian TNF-Rs is suggestive of additional conservative substitutions that may be made without altering the essential biological characteristics of TNF-R or IL-1R.

Alterations of the native amino acid sequence may be accomplished by any of a number of known techniques. Mutations can be introduced at particular loci by

synthesizing oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following ligation, the resulting reconstructed sequence encodes an analog having the desired amino acid insertion, substitution, or deletion.

5           Alternatively, oligonucleotide-directed site-specific mutagenesis procedures can be employed to provide an altered gene having particular codons altered according to the substitution, deletion, or insertion required. Exemplary methods of making the alterations set forth above are disclosed by Walder et al. (*Gene* 42:133, 1986); Bauer et al. (*Gene* 37:73, 1985); Craik (*BioTechniques*, January 1985, 12-19); Smith et al. (*Genetic Engineering: Principles and Methods*, Plenum Press, 1981); and U.S. Patent  
10 Nos. 4,518,584 and 4,737,462.

          The variant amino acid sequence preferably is at least 80% identical, most preferably at least 90% identical, to the native sequence. Percent similarity may be  
15 determined, for example, by comparing sequence information using the GAP computer program, version 6.0, available from the University of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443, 1970), as revised by Smith and Waterman (*Adv. Appl. Math.* 2:482, 1981). Briefly, the GAP program defines similarity as the number  
20 of aligned symbols (i.e., nucleotides or amino acids) which are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred default parameters for the GAP program include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) for nucleotides, and the weighted comparison matrix of Gribskov and Burgess, *Nucl. Acids Res.* 14:6745, 1986, as  
25 described by Schwartz and Dayhoff, eds., *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, pp. 353-358, 1979; (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

          Other proteins capable of blocking the binding of IL-1 to cellular receptors *in vivo* may be substituted for the IL-1R polypeptides described above in the fusion  
30 proteins of the present invention. Such proteins, generally referred to as IL-1 receptor antagonists, include those described by Eisenberg et al. (*Nature* 343:341, 1990), Hannum et al. (*Nature* 343:336, 1990), and Carter et al. (*Nature* 344:633, 1990). These antagonist proteins bind to IL-1 receptors, but have no  
35 IL-1-like activity (e.g., do not transduce a signal or otherwise produce the biological effects that result from binding of IL-1 to a

cellular IL-1 receptor). The antagonist proteins compete with IL-1 for binding to endogenous IL-1 receptors, thus inhibiting biological effects mediated by IL-1 *in vivo*.

#### DNA Sequences Encoding Recombinant Fusion Proteins

5 Isolated DNA sequences encoding the above-described fusion proteins are also provided by the present invention. A DNA sequence encoding a fusion protein of the present invention is constructed using recombinant DNA techniques to insert DNA fragments encoding the IL-1R or TNF-R polypeptides into an appropriate expression vector. The 3' end of a DNA fragment encoding TNF-R is ligated (*via* a peptide linker) 10 to the 5' end of the DNA fragment encoding IL-1R with the reading frames of the sequences in phase to permit translation of the mRNA into a single biologically active fusion protein. Alternatively, the 3' end of a DNA fragment encoding IL-1R may be ligated (*via* a peptide linker) to the 5' end of the DNA fragment encoding TNF-R, with the reading frames of the sequences in phase to permit translation of the mRNA into a 15 single biologically active fusion protein. An additional sequence encoding TNF-R may be ligated in the same reading frame to produce a sequence encoding a fusion protein comprising two TNF-R polypeptides and one IL-1R polypeptide. The IL-1R-encoding sequence is preferably positioned upstream of the TNF-R-encoding sequence(s). While the fusion protein may comprise two IL-1R polypeptides along with the TNF-R 20 polypeptide(s), one IL-1R is preferred. A single IL-1R provides the desired high affinity IL-1 binding activity without the possible disadvantages of increasing the size of the fusion protein by adding a second IL-1R. Such disadvantages may include increased complexity of vector construction procedures and possible reduction in the level of expression of the desired protein. In another embodiment of the present 25 invention, a DNA sequence encoding TNF-R is ligated to a linker sequence which in turn is ligated to a second TNF-R encoding sequence.

A DNA sequence encoding an N-terminal signal sequence may be retained on the DNA sequence encoding the N-terminal polypeptide, while stop codons, which would prevent read-through to the downstream DNA sequence(s), are eliminated. 30 Conversely, a stop codon required to end translation is generally retained on the DNA sequence encoding the C-terminal polypeptide. DNA encoding a signal sequence is preferably removed from DNA sequences other than those encoding the N-terminal polypeptide.

A DNA sequence encoding a desired peptide linker may be inserted between, 35 and in the same reading frame as, the DNA sequences encoding TNF-R or IL-1R using any suitable conventional technique. For example, a chemically synthesized

oligonucleotide encoding the linker and containing appropriate restriction endonuclease cleavage sites may be ligated between the sequences encoding TNF-R or IL-1R. Alternatively, a chemically synthesized DNA sequence may contain a sequence complementary to the 3' terminus (without the stop codon) of either TNF-R or IL-1R followed by a linker-encoding sequence which is followed by a sequence complementary to the 5' terminus of the other of TNF-R and IL-1R. Oligonucleotide directed mutagenesis is then employed to insert the linker-encoding sequence into a vector containing a direct fusion of TNF-R and IL-1R. Another technique employs polymerase chain reactions using primers comprising, in part, single strand segments encoding a peptide linker. PCR-generated DNA fragments encoding two different proteins can be joined through annealing of the complementary single stranded linker-encoding segments present at a terminus of each fragment. Preferred procedures for inserting a linker-encoding DNA segment between TNF-R and IL-1R DNA sequences (or between two TNF-R DNA sequences) are described in examples 11 and 12 below.

DNA sequences encoding TNF-R and IL-1R may be isolated by any suitable conventional procedure, for use in constructing the fusion protein-encoding DNA sequences of the present invention. DNA sequences encoding fusion proteins to be expressed in a microorganism will preferably contain no introns that could prematurely terminate transcription of DNA into mRNA; however, premature termination of transcription may be desirable, for example, where it would result in mutants having advantageous C-terminal truncations, for example, deletion of a transmembrane region to yield a soluble receptor not bound to the cell membrane. Among the suitable procedures for cloning type I and type II IL-1R cDNA are those presented in examples 8 and 9, respectively. TNF-R cDNA may be isolated by procedures that include those described in example 3.

The coding sequence of TNF-R may be obtained by isolating a sequence encoding TNF-R from a recombinant cDNA or genomic DNA library. A cDNA library is preferably constructed by obtaining polyadenylated mRNA from a particular cell line which expresses a mammalian TNF-R, for example, the human fibroblast cell line WI-26 VA4 (ATCC CCL 95.1) and using the mRNA as a template for synthesizing double stranded cDNA. The double stranded cDNA is then packaged into a recombinant vector, which is introduced into a host cell (e.g., an appropriate *E. coli* strain) and propagated. TNF-R sequences contained in the cDNA library can be identified by screening the library with an appropriate nucleic acid probe which is capable of hybridizing with human TNF-R cDNA. Another cloning technique that may be employed is the direct expression procedure described in example 3 below.

Alternatively, DNAs encoding TNF-R proteins can be assembled by ligation of synthetic oligonucleotide subunits corresponding to all or part of the sequence of Figure 2 to provide a complete coding sequence.

5 Additional cDNA clones can be isolated from cDNA libraries of other mammalian species by cross-species hybridization. For use in hybridization, DNA encoding TNF-R or IL-1R may be covalently labeled with a detectable substance such as a fluorescent group, a radioactive atom or a chemiluminescent group by methods well known to those skilled in the pertinent art.

10 Like most mammalian genes, mammalian TNF receptors and IL-1 receptors are presumably encoded by multi-exon genes. Alternative mRNA constructs which can be attributed to different mRNA splicing events following transcription, and which share large regions of identity or similarity with the cDNAs disclosed herein such that biologically active TNF-R or IL-1R is encoded thereby, are considered to be useful in preparing the fusion proteins of the present invention.

15 DNA encoding soluble TNF-R and IL-1R polypeptides may be prepared by any of a number of conventional techniques. A DNA fragment encoding a desired soluble polypeptide may be subcloned into an expression vector. DNA fragments may be produced by restriction endonuclease digestion of a full length cloned DNA sequence, and isolated by electrophoresis on agarose gels. Alternatively, a desired DNA sequence  
20 may be chemically synthesized using known techniques. Linkers containing restriction endonuclease cleavage site(s) may be employed to insert the desired DNA fragment into an expression vector, or the fragment may be digested at cleavage sites naturally present therein.

25 The well known polymerase chain reaction (PCR) procedures also may be employed to isolate a DNA sequence encoding a desired soluble protein fragment. This technique is illustrated in the examples below.

30 In another approach, enzymatic treatment (using Bal 31 exonuclease) may be employed to delete terminal nucleotides from a DNA fragment to obtain a fragment having a particular desired terminus. Among the commercially available linkers are those that can be ligated to the blunt ends produced by Bal 31 digestion, and which contain restriction endonuclease cleavage site(s). Alternatively, oligonucleotides that reconstruct the N- or C- terminus of a DNA fragment to a desired point may be synthesized. The oligonucleotide may contain a restriction endonuclease cleavage site upstream of the desired coding sequence and position an initiation codon (ATG) at the  
35 N-terminus of the coding sequence.

The TNF-R and IL-1R DNA sequences may vary from those presented in SEQ ID NOS:1, 3, 5 and 7. Due to the known degeneracy of the genetic code, there can be considerable variation in nucleotide sequences encoding the same amino acid sequence, for example. DNA sequences capable of hybridizing to the DNA sequences of SEQ ID NOS:1, 3, 5 and 7 under moderately stringent conditions (55°C, 5 X SSC), and which encode a biologically active TNF-R or IL-1R polypeptide, are also considered to be TNF-R-encoding or IL-1R-encoding DNA sequences, respectively, in the context of the present invention. Mutations may be deliberately made to the native DNA sequences, e.g. to produce the amino acid substitutions, deletions and insertions described above. Certain of the mutations will not be expressed in the final protein product. For example, nucleotide substitutions may be made to enhance expression, primarily to avoid secondary structure loops in the transcribed mRNA (see EP 75,444). Other alterations of the nucleotide sequence may be made to provide codons that are more readily translated by the selected host, e.g., the well-known *E. coli* preference codons for *E. coli* expression. Silent mutations (changes in the DNA sequence that do not alter the encoded amino acid sequence) also may occur during polymerase chain reactions. In one embodiment of the present invention, nucleotide number 437 of SEQ ID NO:5 (type I IL-1R) is changed from a T to a C. This silent mutation occurred during a polymerase chain reaction.

Mutations in nucleotide sequences should, of course, preserve the reading frame phase of the coding sequences. The mutations preferably will not create complementary regions that could hybridize to produce secondary mRNA structures such as loops or hairpins which would adversely affect translation of the fusion protein mRNA.

The present invention thus provides inventive DNA sequences encoding the above-described fusion proteins, wherein each TNF-R DNA sequence in the fusion protein DNA sequence is selected from: (a) DNA sequences derived from the coding region of a native mammalian TNF-R gene (e.g., cDNA derived from the coding region of SEQ ID NOS:1 or 3); (b) DNA sequences capable of hybridization to a DNA sequence of (a) under moderately stringent conditions (50°C, 2x SSC) and which encode biologically active TNF-R and (c) DNA sequences which are degenerate as a result of the genetic code to the DNA sequences defined in (a) or (b) and which encode biologically active TNF-R. The fusion protein DNA sequence likewise may comprise IL-1R encoding DNA sequence(s) selected from: (a) DNA sequences derived from the coding region of a native mammalian IL-1R gene (e.g., cDNA derived from the coding region of SEQ ID NOS:5 or 7); (b) DNA sequences capable of hybridization to a DNA

sequence of (a) under moderately stringent conditions (55°C, 5x SSC) and which encode biologically active IL-1R; and (c) DNA sequences which are degenerate as a result of the genetic code to the DNA sequences defined in (a) or (b) and which encode biologically active IL-1R.

5

#### Expression of Recombinant Fusion Proteins

The present invention provides recombinant expression vectors to express DNA encoding the fusion proteins of the present invention. The inventive recombinant expression vectors are replicable DNA constructs which contain a synthetic or cDNA-  
10 derived DNA sequence encoding one of the above-described fusion proteins, operably linked to suitable transcriptional or translational regulatory elements. Examples of genetic elements having a regulatory role in gene expression include transcriptional promoters, operators or enhancers, a sequence encoding suitable mRNA ribosomal binding sites, and appropriate transcription and translation initiation and termination  
15 sequences. The ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants may additionally be incorporated. The regulatory elements employed in the expression vectors are generally derived from mammalian, microbial, viral, or insect genes. Expression vectors derived from retroviruses also may be employed.

20 DNA regions are operably linked when they are functionally related to each other. A DNA sequence encoding a fusion protein is said to be operably linked to one or more of the above-described regulatory elements when the fusion protein DNA sequence is transcribed, or the resulting mRNA is translated, under the control of the regulatory element(s).

25 Transformed host cells are cells which have been transformed or transfected with foreign DNA using recombinant DNA techniques. In the context of the present invention, the foreign DNA includes a sequence encoding the inventive fusion protein. Host cells may be transformed for purposes of cloning or amplifying the foreign DNA, or may be transformed with an expression vector for production of the fusion protein  
30 under the control of appropriate promoters. Suitable host cells include prokaryotes, yeast, or higher eukaryotic cells. Appropriate cloning and expression vectors for use with bacterial, fungal, yeast, and mammalian cellular hosts are described by Pouwels et al. (*Cloning Vectors: A Laboratory Manual*, Elsevier, New York, 1985).

#### Cell-free translation systems

35 could also be employed to produce fusion protein using RNAs derived from the DNA constructs of the present invention.

Prokaryotes include gram negative or gram positive organisms. Prokaryotic expression vectors generally comprise one or more phenotypic selectable markers, for example a gene encoding proteins conferring antibiotic resistance or supplying an autotrophic requirement, and an origin of replication recognized by the host to ensure  
5 amplification within the host. Examples of suitable prokaryotic hosts for transformation include *E. coli*, bacilli such as *Bacillus subtilis*, *Salmonella typhimurium*, and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

Useful expression vectors for bacterial use can comprise a selectable marker and  
10 bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well-known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and pGEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the  
15 structural sequence to be expressed. *E. coli* is typically transformed using derivatives of pBR322, a plasmid derived from an *E. coli* species (Bolivar et al., *Gene* 2:95, 1977). pBR322 contains genes for ampicillin and tetracycline resistance, providing simple means for identifying transformed cells.

Promoters commonly used in recombinant microbial expression vectors include  
20 the b-lactamase (penicillinase) and lactose promoter system (Chang et al., *Nature* 275:615, 1978; and Goeddel et al., *Nature* 281:544, 1979), the tryptophan (*trp*) promoter system (Goeddel et al., *Nucl. Acids Res.* 8:4057, 1980; and EPA 36,776) and *tac* promoter (Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, p. 412, 1982). A particularly useful bacterial expression system  
25 employs the phage  $\lambda$  P<sub>L</sub> promoter and cI857ts thermoinducible repressor. Plasmid vectors available from the American Type Culture Collection which incorporate derivatives of the  $\lambda$  P<sub>L</sub> promoter include plasmid pHUB2, resident in *E. coli* strain JMB9 (ATCC 37092) and pPLc28, resident in *E. coli* RR1 (ATCC 53082).

The recombinant fusion protein may also be expressed in yeast hosts,  
30 preferably from *Saccharomyces* species, such as *S. cerevisiae*. Yeast of other genera such as *Pichia* or *Kluyveromyces* may also be employed. Yeast vectors will generally contain an origin of replication from the 2 $\mu$ m yeast plasmid or an autonomously replicating sequence (ARS), a promoter, DNA encoding the fusion protein, sequences for polyadenylation and transcription termination and a selection gene. Preferably,  
35 yeast vectors will include an origin of replication and selectable markers permitting transformation of both yeast and *E. coli*, e.g., the ampicillin resistance gene of *E. coli*



and the *S. cerevisiae* *trp1* gene, which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, and a promoter derived from a highly expressed yeast gene to induce transcription of a structural sequence downstream. The presence of the *trp1* lesion in the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoter sequences in yeast vectors include the promoters for metallothionein, 3-phosphoglycerate kinase (Hitzeman et al., *J. Biol. Chem.* 255:2073, 1980) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Reg.* 7:149, 1968; and Holland et al., *Biochem.* 17:4900, 1978), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase and glucokinase. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., EPA 73,657.

Preferred yeast vectors can be assembled using DNA sequences from pBR322 for selection and replication in *E. coli* (*Amp<sup>r</sup>* gene and origin of replication) and yeast DNA sequences including a glucose-repressible ADH2 promoter and  $\alpha$ -factor secretion leader. The ADH2 promoter has been described by Russell et al. (*J. Biol. Chem.* 258:2674, 1982) and Beier et al., (*Nature* 300:724, 1982). Advantageously, a DNA segment encoding a leader sequence functional in yeast is operably linked to the 5' end of the DNA encoding the fusion protein. The encoded leader peptide promotes secretion of the fusion protein from the host cell and is generally cleaved from the fusion protein upon secretion. As one example, the yeast  $\alpha$ -factor leader, which directs secretion of heterologous proteins, can be inserted between the promoter and the structural gene to be expressed. See, e.g., Kurjan et al., *Cell* 30:922, 1982; and Bitter et al., *Proc. Natl. Acad. Sci. USA* 81:5330, 1984. The leader sequence may be modified to contain, near its 3' end, one or more useful restriction sites to facilitate fusion of the leader sequence to foreign genes.

Suitable yeast transformation protocols are known to those of skill in the art. An exemplary technique is described by Hinnen et al., *Proc. Natl. Acad. Sci. USA* 75:1929, (1978), selecting for *Trp<sup>+</sup>* transformants in a selective medium consisting of 0.67% yeast nitrogen base, 0.5% casamino acids, 2% glucose, 10 $\mu$ g/ml adenine and 20  $\mu$ g/ml uracil. Host strains transformed by vectors comprising the above-described ADH2 promoter may be grown for expression in a rich medium consisting of 1% yeast extract, 2% peptone, and 1% glucose supplemented with 80  $\mu$ g/ml adenine and 80  $\mu$ g/ml uracil. Derepression of the ADH2 promoter occurs upon exhaustion of medium

glucose. Crude yeast supernatants are harvested by filtration and held at 4°C prior to further purification.

Various mammalian or insect cell culture systems can be employed to express recombinant protein. Baculovirus systems for production of heterologous proteins in  
5 insect cells are reviewed by Luckow and Summers, *Bio/Technology* 6:47 (1988). Established cell lines of mammalian origin may be employed. Examples of suitable mammalian host cell lines include the COS-7 lines of monkey kidney cells, described by Gluzman (*Cell* 23:175, 1981), L cells, C127, 3T3, Chinese hamster ovary (CHO), HeLa and BHK cell lines. Mammalian expression vectors may comprise non-  
10 transcribed elements such as an origin of replication, a suitable promoter and enhancer linked to the gene to be expressed, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences, such as necessary ribosome binding sites, a poly-adenylation site, splice donor and acceptor sites, and transcriptional termination sequences.

15 The transcriptional and translational control sequences in expression vectors to be used in transforming vertebrate cells may be provided by viral sources. For example, commonly used promoters and enhancers are derived from Polyoma, Adenovirus 2, Simian Virus 40 (SV40), and human cytomegalovirus. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early and late  
20 promoter, enhancer, splice, and polyadenylation sites may be used to provide the other genetic elements required for expression of a heterologous DNA sequence. The early and late promoters are particularly useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin or replication (Fiers et al., *Nature* 273:113, 1978). Smaller or larger SV40 fragments may also be used, provided  
25 the approximately 250 bp sequence extending from the Hind III site toward the *Bgl*I site located in the viral origin of replication is included. Exemplary vectors can be constructed as disclosed by Okayama and Berg (*Mol. Cell. Biol.* 3:280, 1983). A useful system for stable high level expression of mammalian receptor cDNAs in C127 murine mammary epithelial cells can be constructed substantially as described by  
30 Cosman et al. (*Mol. Immunol.* 23:935, 1986).

#### Producing and Purifying the Fusion Protein

The present invention provides a process for producing the recombinant fusion protein of the present invention, comprising culturing a host cell transformed with an  
35 expression vector comprising a DNA sequence that encodes said fusion protein under conditions that promote expression of the fusion protein, which is then purified from

5 culture media or cell extracts. Any suitable purification process may be employed, with the procedure of choice varying according to such factors as the type of host cells and whether or not the desired protein is secreted from the host cells. The fusion protein will be secreted into the culture medium when it is initially fused to a signal sequence or leader peptide operative in the host cells, or when the protein comprises soluble forms of the TNF-R and IL-1R polypeptides.

10 For example, supernatants from expression systems which secrete recombinant protein into the culture medium can be first concentrated using a commercially available protein concentration filter, for example, an Amicon\* or Millipore Pellicon\* ultrafiltration unit. Following the concentration step, the concentrate can be applied to a suitable purification matrix. For example, a suitable affinity matrix can comprise TNF or IL-1. An affinity matrix may be prepared by coupling recombinant human TNF or IL-1 to cyanogen bromide-activated Sepharose\* (Pharmacia) or Hydrazide Affigel\* (Biorad), according to manufacturer's recommendations. A preferred purification procedure  
15 involves sequential immunopurification using antibodies bound to a suitable support. Proteins binding to an antibody specific for TNF-R are recovered and contacted with antibody specific for IL-1R on an insoluble support. Proteins immunoreactive with both antibodies may thus be identified and isolated. A monoclonal antibody specific for human type I IL-1R was deposited with the American Type Culture Collection under  
20 accession number HB10556 on September 13, 1990. Alternatively, an anion exchange resin can be employed, for example, a matrix or substrate having pendant diethylaminoethyl (DEAE) groups. The matrices can be acrylamide, agarose, dextran, cellulose or other types commonly employed in protein purification. Alternatively, a cation exchange step can be employed. Suitable cation exchangers include various  
25 insoluble matrices comprising sulfopropyl or carboxymethyl groups. Sulfopropyl groups are preferred. One or more reversed-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify a fusion protein composition.

30 Recombinant protein produced in bacterial culture is usually isolated by initial extraction from cell pellets, followed by one or more concentration, salting-out, aqueous ion exchange or size exclusion chromatography steps. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of recombinant fusion proteins can  
35 disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

\*Trade-mark

Fermentation of yeast which express fusion proteins as a secreted protein greatly simplifies purification. Secreted recombinant protein resulting from a large-scale fermentation can be purified by methods analogous to those disclosed by Urdal et al. (*J. Chromatog.* 296:171, 1984), involving two sequential, reversed-phase HPLC steps for purification of a recombinant protein on a preparative HPLC column.

Some or all of the foregoing purification steps, in various combinations, can be employed to provide an essentially homogeneous recombinant protein. Recombinant cell culture enables the production of the fusion protein free of those contaminating proteins which may be normally associated with TNF-R or IL-1R as they are found in nature in their respective species of origin, e.g., in cells, cell exudates or body fluids. The foregoing purification procedures are among those that may be employed to purify non-recombinant receptors of the present invention as well.

As an alternative to production of the inventive receptors as fusion proteins, the TNF-R and IL-1R proteins may be separately produced and purified, and subsequently linked together. Numerous reagents useful for crosslinking one protein molecule to another are known. Heterobifunctional and homobifunctional linkers are available for this purpose from Pierce Chemical Company, Rockford, Illinois, for example. Such linkers contain two functional groups (e.g., esters and/or maleimides) that will react with certain functional groups on amino acid side chains (e.g., amines on lysine residues and sulfhydryls generated on cysteine residues by reduction), thus linking one polypeptide to another. Examples of such crosslinking reagents are N-maleimidobenzoyl succinimidyl ester and N-hydroxysuccinimide. The reagent and reaction conditions should be chosen such that the cross-linking does not interfere with binding of TNF-R or IL-1 to the receptor. The TNF-R and IL-1R polypeptides are preferably linked *via* one of the above-described peptide linkers that functions as a spacer. A peptide linker may be attached to TNF-R or to IL-1R by any of the conventional procedures used to attach one polypeptide to another. The cross-linking reagents available from Pierce Chemical Company as described above are among those that may be employed. Amino acids having side chains reactive with such reagents may be included in the peptide linker, e.g., at the termini thereof.

#### Pharmaceutical Compositions

The present invention provides pharmaceutical compositions comprising any of the above-described fusion proteins and a physiologically acceptable carrier, diluent, or excipient. Such carriers, excipients and diluents will be nontoxic to recipients at the dosages and concentrations employed. Such compositions may comprise buffers,

antioxidants such as ascorbic acid, low molecular weight (less than about 10 residues) polypeptides, proteins, amino acids, carbohydrates including glucose, sucrose or dextrans, chelating agents such as EDTA, glutathione and other stabilizers and excipients. Neutral buffered saline or saline mixed with conspecific serum albumin are  
5 exemplary appropriate diluents. Preferably, the composition is formulated as a lyophilizate using appropriate excipient solutions (e.g., sucrose) as diluents. Appropriate dosages can be determined in clinical trials. The amount and frequency of administration will depend, of course, on such factors as the nature and severity of the indication being treated, the desired response, the condition of the patient, and so forth.

10 Conditions mediated by either TNF or IL-1 may be treated by administering a therapeutically effective amount of a fusion protein of the present invention in the form of a pharmaceutical composition, to a patient afflicted with such a disorder. A disorder is said to be mediated by TNF or IL-1 when TNF or IL-1 causes (directly or indirectly) or exacerbates the disorder. Soluble receptor proteins can be used to competitively bind  
15 to TNF or IL-1, thereby inhibiting binding of TNF and IL-1 to cell surface receptors.

For therapeutic use, purified fusion proteins of the present invention are administered to a patient, preferably a human, for treatment in a manner appropriate to the indication. Thus, for example, the pharmaceutical compositions can be administered by bolus injection, continuous infusion, sustained release from implants,  
20 or other suitable technique.

The fusion protein employed in the pharmaceutical compositions should be purified, in that the fusion protein is substantially free of other proteins of natural or endogenous origin and contains less than about 1% by mass of protein contaminants residual of production processes. Such compositions, however, can contain other  
25 proteins added as stabilizers, carriers, excipients or co-therapeutics. The fusion protein is purified to substantial homogeneity if it is detectable as a single protein band in a polyacrylamide gel by silver staining.

The fusion proteins of the present invention may be administered to treat conditions believed to be mediated, at least in part, by TNF, such as cachexia,  
30 rheumatoid arthritis, diabetes, multiple sclerosis, pulmonary fibrosis and silicosis, cerebral malaria, and allograft and xenograft rejection in graft versus host disease. TNF has also been implicated in sepsis and septic shock. Bacterial endotoxin can cause sepsis in mammals infected with certain types of bacteria, and is believed to stimulate macrophages to produce factors that include TNF. Folks et al. (*PNAS USA* 86:2365,  
35 1989) suggests that TNF- $\alpha$  plays an important role in the pathogenesis of HIV

infection. TNF- $\alpha$  induced expression of HIV in a cell line employed as a model of HIV latency to study conversion from a latent to a productive infection.

Certain cytokines (IL-1, IL-2 and other colony stimulating factors) can induce significant host production of TNF. Fusion proteins of the formula TNF-R-linker-  
5 TNF-R thus may be used to treat side effects associated with cytokine therapy. Because of the primary role IL-1 plays in the production of TNF, fusion proteins comprising both IL-1 receptor(s) and TNF receptor(s) may be preferred in the treatment of TNF-associated clinical indications.

TNF has been reported to induce secretion of IL-1 *in vivo*. Thus, fusion  
10 proteins that bind both TNF and IL-1 may be employed in treating conditions mediated by IL-1. The inventive fusion proteins can be administered, for example, for the purpose of suppressing immune responses in a human. A variety of diseases or conditions are caused by an immune response to alloantigen. In alloantigen-induced immune responses, IL-1R suppresses lymphoproliferation and inflammation which  
15 result upon activation of T cells. IL-1R thus may be used to suppress alloantigen-induced immune responses in the clinical treatment of, for example, rejection of allografts (such as skin, kidney, and heart transplants), and graft-versus-host reactions in patients who have received bone marrow transplants. IL-1 is believed to play a causative role in allergies and autoimmune dysfunctions (such as rheumatoid arthritis,  
20 diabetes, and multiple sclerosis, which are dependent upon the activation of T cells against antigens not recognized as being indigenous to the host.

TNF and IL-1 have been implicated in a number of the same diseases, as can be seen by comparing the lists of TNF-mediated and IL-1-mediated conditions presented above. In addition, TNF and IL-1 are two of the major mediators of inflammation,  
25 often acting in concert. The fusion proteins of the present invention that comprise receptors for both TNF and IL-1 thus offer advantages in the treatment of a number of conditions in which both TNF and IL-1 are believed to play a causative role.

The use of soluble forms of IL-1R and TNF-R in the inventive fusion proteins is advantageous for certain applications. Purification of the proteins from recombinant  
30 host cells is facilitated, since the soluble proteins are secreted from the cells. Further, soluble proteins are generally more suitable for intravenous administration and may exert their therapeutic effect (binding IL-1 and/or TNF) in the bloodstream. By binding IL-1 and/or TNF, the soluble fusion proteins will inhibit signal transduction via endogenous cell surface receptors for IL-1 or TNF.

The inventive fusion proteins may also be used as reagents in receptor-based immunoassays, reagents in assays for TNF or IL-1, or as binding agents for affinity purification or TNF or IL-1.

5 The following examples are offered by way of illustration, and not by way of limitation.

## EXAMPLES

10

### Example 1

#### TNF Binding Assays

15

20

25

A. *Radiolabeling of TNF $\alpha$  and TNF $\beta$* . Recombinant human TNF $\alpha$ , in the form of a fusion protein containing a hydrophilic octapeptide at the N-terminus, was expressed in yeast as a secreted protein and purified by affinity chromatography (Hopp et al., *Bio/Technology* 6:1204, 1988). Purified recombinant human TNF $\beta$  was purchased from R&D Systems (Minneapolis, MN). Both proteins were radiolabeled using the commercially available solid phase agent, IODO-GEN\* (Pierce). In this procedure, 5  $\mu$ g of IODO-GEN\* were plated at the bottom of a 10 x 75 mm glass tube and incubated for 20 minutes at 4°C with 75  $\mu$ l of 0.1 M sodium phosphate, pH 7.4 and 20  $\mu$ l (2 mCi) Na <sup>125</sup>I. This solution was then transferred to a second glass tube containing 5  $\mu$ g TNF $\alpha$  (or TNF $\beta$ ) in 45  $\mu$ l PBS for 20 minutes at 4°C. The reaction mixture was fractionated by gel filtration on a 2 ml bed volume of Sephadex\* G-25 (Sigma) equilibrated in Roswell Park Memorial Institute (RPMI) 1640 medium containing 2.5% (w/v) bovine serum albumin (BSA), 0.2% (w/v) sodium azide and 20 mM Hepes pH 7.4 (binding medium). The final pool of <sup>125</sup>I-TNF was diluted to a working stock solution of 1 x 10<sup>-7</sup> M in binding medium and stored for up to one month at 4°C without detectable loss of receptor binding activity. The specific activity is routinely 1 x 10<sup>6</sup> cpm/mmol TNF.

30

35

B. *Binding to Intact Cells*. Binding assays with intact cells were performed by two methods. In the first method, cells were first grown either in suspension (e.g., U 937) or by adherence on tissue culture plates (e.g., WI26-VA4 or COS cells expressing the recombinant TNF receptor). Adherent cells were subsequently removed by treatment with 5mM EDTA treatment for ten minutes at 37 degrees centigrade. Binding assays were then performed by a phalate oil separation method (Dower et al., *J. Immunol.* 132:751, 1984) essentially as described by Park et al. (*J. Biol. Chem.* 261:4177, 1986). Non-specific binding of <sup>125</sup>I-TNF was measured in the presence of

\*Trade-mark

a 200-fold or greater molar excess of unlabeled TNF. Sodium azide (0.2%) was included in a binding assay to inhibit internalization of  $^{125}\text{I}$ -TNF by cells. In the second method, COS cells transfected with the TNF-R-containing plasmid, and expressing TNF receptors on the surface, were tested for the ability to bind  $^{125}\text{I}$ -TNF by the plate binding assay described by Sims et al. (*Science* 241:585, 1988).

5  
10  
15  
20  
C. *Solid Phase Binding Assays*. The ability of TNF-R to be stably adsorbed to nitrocellulose from detergent extracts of human cells yet retain TNF-binding activity provided a means of detecting TNF-R. Cell extracts were prepared by mixing a cell pellet with a 2 x volume of PBS containing 1% Triton<sup>\*</sup> X-100 and a cocktail of protease inhibitors (2 mM phenylmethyl sulfonyl fluoride, 10  $\mu\text{M}$  pepstatin, 10  $\mu\text{M}$  leupeptin, 2 mM o-phenanthroline and 2 mM EGTA) by vigorous vortexing. The mixture was incubated on ice for 30 minutes after which it was centrifuged at 12,000x g for 15 minutes at 8°C to remove nuclei and other debris. Two microliter aliquots of cell extracts were placed on dry BA85/21 nitrocellulose membranes (Schleicher and Schuell, Keene, NH) and allowed to dry. The membranes were incubated in tissue culture dishes for 30 minutes in Tris (0.05 M) buffered saline (0.15 M) pH 7.5 containing 3% w/v BSA to block nonspecific binding sites. The membrane was then covered with  $5 \times 10^{-11}$  M  $^{125}\text{I}$ -TNF in PBS + 3% BSA and incubated for 2 hr at 4°C with shaking. At the end of this time, the membranes were washed 3 times in PBS, dried and placed on Kodak X-Omat<sup>\*</sup> AR film for 18 hr at -70°C.

25  
30  
35  
D. *Signal Transduction Assays*. Inhibition of TNF signal transduction activity can be determined by transfecting cells with recombinant TNF-R DNAs encoding membrane-bound TNF-R to obtain recombinant receptor expression on the cell surface. The cells are then contacted with TNF and the resulting metabolic effects examined. If an effect results which is attributable to the action of the ligand, and is not attributable to endogenous TNF receptors on the cells, then the recombinant receptor has signal transduction activity. Exemplary procedures for determining whether a polypeptide has signal transduction activity are disclosed by Idzerda et al., *J. Exp. Med.* 171:861 (1990); Curtis et al., *Proc. Natl. Acad. Sci. USA* 86:3045 (1989); Prywes et al., *EMBO J.* 5:2179 (1986) and Chou et al., *J. Biol. Chem.* 262:1842 (1987). The ability of a soluble TNF-R polypeptide to competitively inhibit signal transduction can be determined using similar procedures. Primary cells or cell lines which express an endogenous TNF receptor and have a detectable biological response to TNF could be utilized as an alternative to the cells expressing recombinant membrane-bound TNF-R. Decreased signal transduction when a soluble TNF-R

\*Trade-mark



polypeptide is added to the assay indicates binding of TNF by the soluble TNF-R, so that less TNF binds to the cell surface TNF receptors to initiate signal transduction.

5

### Example 2.

#### IL-1 Binding Assays

A. *Radiolabeling of rIL-1 $\beta$* . Recombinant human IL-1 $\beta$  was prepared by expression in *E. coli* and purification to homogeneity as described by Kronheim et al. (10) (*Bio/Technology* 4:1078, 1986). The IL-1 $\beta$  was labeled with di-iodo ( $^{125}\text{I}$ ) Bolton-Hunter reagent (New England Nuclear, Glenolden, PA). Ten micrograms (0.57 nmol) of protein in 10  $\mu\text{L}$  of phosphate (0.015 mol/L)-buffered saline (PBS; 0.15 mol/L), pH 7.2, was mixed with 10  $\mu\text{L}$  of sodium borate (0.1 mol/L)-buffered saline (0.15 mol/L), pH 8.5, and reacted with 1 mCi (0.23 nmol) of Bolton-Hunter reagent according to the (15) manufacturer's instructions for 12 hours at 8°C. Subsequently, 30  $\mu\text{L}$  of 2% gelatin and 5  $\mu\text{L}$  of 1 mol/L glycine ethyl ester were added, and the protein was separated from unreacted Bolton-Hunter reagent on a 1 mL bed volume Biogel<sup>TM</sup> P6 column (BioRad Laboratories, Richmond, CA). Routinely, 50% to 60% incorporation of label was observed. Radioiodination yielded specific activities in the range of  $1 \times 10^{15}$  to  $5 \times$  (20)  $10^{15}$  cpm/mmol-1 (0.4 to 2 atoms I per molecule protein), and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS/PAGE) revealed a single labeled polypeptide of 17.5 kD, consistent with previously reported values for IL-1. The labeled protein was greater than 98% TCA precipitable, indicating that the  $^{125}\text{I}$  was covalently bound to protein.

25

B. *Inhibition Binding Assay for Membrane-Bound IL-1R*. "IL-1" refers collectively to IL-1 $\alpha$  and IL-1 $\beta$ . The binding inhibition constant of an IL-1R protein may be determined by inhibition binding assays in which varying concentrations of a competitor (IL-1 $\beta$  or IL-1 $\alpha$ ) are incubated with a constant amount of radiolabeled IL-1 $\beta$  or IL-1 $\alpha$  and cells expressing the IL-1R. The non-radiolabeled competitor binds to the (30) receptor and prevents the radiolabeled ligand from binding to the receptor. Binding assays were performed by a phthalate oil separation method essentially as described by Dower et al., *J. Immunol.* 132:751, 1984 and Park et al., *J. Biol. Chem.* 261:4177, 1986. Briefly, host cells expressing a membrane-bound recombinant IL-1R were incubated in six-well plates (Costar, Cambridge, MA) at 4°C for 2 hours with  $^{125}\text{I}$ -IL- (35) 1 $\beta$  in 1 ml binding medium (Roswell Park Memorial Institute (RPMI) 1640 medium combining 2% BSA, 20 mM hepes buffer, and 0.1% sodium azide, pH 7.2). Sodium

azide was included to inhibit internalization and degradation of  $^{125}\text{I}$ -IL-1 by cells at  $37^\circ\text{C}$ . The plates were incubated on a gyratory shaker for 1 hour at  $37^\circ\text{C}$ . Replicate aliquots of the incubation mixture were then transferred to polyethylene centrifuge tubes containing a phthalate oil mixture comprising 1.5 parts dibutylphthalate, to 1 part bis(s-ethylhexyl)phthalate. Control tubes containing a 100X molar excess of unlabeled IL-1 $\beta$  were also included to determine non-specific binding. The cells with bound  $^{125}\text{I}$ -IL-1 were separated from unbound  $^{125}\text{I}$ -IL-1 by centrifugation for 5 minutes at 15,000X g in an Eppendorf Microfuge. The radioactivity associated with the cells was then determined on a gamma counter.

10           C. *Inhibition Binding Assay for Soluble IL-1R*. The binding inhibition constant of a soluble human IL-1R may be determined by an inhibition binding assay in which varying concentrations of an IL-1 $\beta$  competitor are incubated with a constant amount of radiolabeled I-IL-1 $\beta$  and CB23 cells (an Epstein Barr virus transformed cord blood B lymphocyte cell line) expressing the type II IL-1R. A cell line expressing  
15 endogenous type I IL-1 receptors may be substituted for the CB23 cells in assays involving soluble type I IL-1R. Binding assays were performed by a phthalate oil separation method essentially as described by Dower et al., *J. Immunol.* 132:751, 1984 and Park et al., *J. Biol. Chem.* 261:4177, 1986. Briefly, CVI-EBNA (mammalian) cells were transfected with the expression vector pDC406 containing a cDNA encoding  
20 a soluble human type II IL-1R as described in example 10. Supernatants from the cells were harvested 3 days after transfection and serially diluted in binding medium (Roswell Park Memorial Institute (RPMI) 1640 medium containing 2% BSA, 20 mM Hepes buffer, and 0.2% sodium azide, pH 7.2) in 6 well plates to a volume of 50  $\mu\text{l}$ /well. The supernatants were incubated with 50  $\mu\text{l}$  of  $9 \times 10^{-10}$  M  $^{125}\text{I}$ -IL-1 $\beta$  plus  
25  $2.5 \times 10^6$  CB23 cells at  $8^\circ\text{C}$  for 2 hours with agitation. Duplicate 60  $\mu\text{l}$  aliquots of the incubation mixture were then transferred to polyethylene centrifuge types containing a phthalate oil mixture comprising 1.5 parts dibutylphthalate, to 1 part bis(s-ethylhexyl)phthalate. A negative control tube containing  $3 \times 10^{-6}$  M unlabeled IL-1 $\beta$  was also included to determine non-specific binding (100% inhibition) and a positive  
30 control tube containing 50 ml binding medium with only radiolabeled IL-1 $\beta$  was included to determine maximum binding. The cells with bound  $^{125}\text{I}$ -IL-1 $\beta$  were separated from unbound  $^{125}\text{I}$ -IL-1 $\beta$  by centrifugation for 5 minutes at 15,000 X g in an Eppendorf Microfuge. Supernatants containing unbound  $^{125}\text{I}$ -IL-1 $\beta$  were discarded and the cells were carefully rinsed with ice-cold binding medium. The cells were  
35 incubated in 1 ml of trypsin-EDTA at  $37^\circ\text{C}$  for 15 minutes and then harvested. The radioactivity associated with the cells was then determined on a gamma counter. The

ability of soluble IL-1R to inhibit binding of IL-1 $\alpha$  to endogenous cellular receptors may be determined by the same procedure. Analogous techniques may be employed in assays involving soluble TNF-R.

5

### Example 3

#### Isolation of Human TNF-R cDNA by Direct Expression of Active Protein in COS-7

##### Cells

10 Various human cell lines were screened for expression of TNF-R based on their ability to bind  $^{125}\text{I}$ -labeled TNF. The human fibroblast cell line WI-26 VA4 (ATCC CCL 95.1) was found to express a reasonable number of receptors per cell. Equilibrium binding studies showed that the cell line exhibited biphasic binding of  $^{125}\text{I}$ -TNF with approximately 4,000 high affinity sites ( $K_a = 1 \times 10^{10} \text{ M}^{-1}$ ) and 15,000 low affinity sites ( $K_a = 1 \times 10^8 \text{ M}^{-1}$ ) per cell.

15 An unsized cDNA library was constructed by reverse transcription of polyadenylated mRNA isolated from total RNA extracted from human fibroblast WI-26 VA4 cells grown in the presence of pokeweed mitogen using standard techniques (Gubler, et al., *Gene* 25:263, 1983; Ausubel et al., eds., *Current Protocols in Molecular Biology*, Vol. 1, 1987). The cells were harvested by lysing the cells in a  
20 guanidine hydrochloride solution and total RNA isolated as previously described (March et al., *Nature* 315:641, 1985).

Poly A<sup>+</sup> RNA was isolated by oligo dT cellulose chromatography and double-stranded cDNA was prepared by a method similar to that of Gubler and Hoffman (*Gene* 25:263, 1983). Briefly, the poly A<sup>+</sup> RNA was converted to an RNA-cDNA hybrid by  
25 reverse transcriptase using oligo dT as a primer. The RNA-cDNA hybrid was then converted into double-stranded cDNA using RNAase H in combination with DNA polymerase I. The resulting double stranded cDNA was blunt-ended with T4 DNA polymerase. To the blunt-ended cDNA is added *EcoRI* linker-adapters (having internal *NotI* sites) which were phosphorylated on only one end (Invitrogen). The linker-  
30 adaptered cDNA was treated with T4 polynucleotide kinase to phosphorylate the 5' overhanging region of the linker-adapter and unligated linkers were removed by running the cDNA over a Sepharose CL4B column. The linker-adaptered cDNA was ligated to an equimolar concentration of *EcoRI* cut and dephosphorylated arms of bacteriophage  $\lambda$ gt10 (Huynh et al, *DNA Cloning: A Practical Approach*, Glover, ed.,  
35 IRL Press, pp. 49-78). The ligated DNA was packaged into phage particles using a commercially available kit to generate a library of recombinants (Stratagene Cloning

Systems, San Diego, CA, USA). Recombinants were further amplified by plating phage on a bacterial lawn of *E. coli* strain c600(hfl<sup>-</sup>).

5 Phage DNA was purified from the resulting  $\lambda$ gt10 cDNA library and the cDNA inserts excised by digestion with the restriction enzyme *Not*1. Following electrophoresis of the digest through an agarose gel, cDNAs greater than 2,000 bp were isolated.

10 The resulting cDNAs were ligated into the eukaryotic expression vector pCAV/NOT, which was designed to express cDNA sequences inserted at its multiple cloning site when transfected into mammalian cells. pCAV/NOT was assembled from pDC201 (a derivative of pMLSV, previously described by Cosman et al., *Nature* 312: 768, 1984), SV40 and cytomegalovirus DNA and comprises, in sequence with the direction of transcription from the origin of replication: (1) SV40 sequences from coordinates 5171-270 including the origin of replication, enhancer sequences and early and late promoters; (2) cytomegalovirus sequences including the promoter and enhancer regions (nucleotides 671 to +63 from the sequence published by Boechart et al. (*Cell* 41:521, 1985); (3) adenovirus-2 sequences containing the first exon and part of the intron between the first and second exons of the tripartite leader, the second exon and part of the third exon of the tripartite leader and a multiple cloning site (MCS) containing sites for *Xho*1, *Kpn*1, *Sma*1, *Not*1 and *Bgl*1; (4) SV40 sequences from coordinates 4127-4100 and 2770-2533 that include the polyadenylation and termination signals for early transcription; (5) sequences derived from pBR322 and virus-associated sequences VAI and VAII of pDC201, with adenovirus sequences 10532-11156 containing the VAI and VAII genes, followed by pBR322 sequences from 4363-2486 and 1094-375 containing the ampicillin resistance gene and origin of replication. pCAV/NOT has been deposited with the American Type Culture Collection under accession no. ATCC 68014.

20 The resulting WI-26 VA4 cDNA library in pCAV/NOT was used to transform *E. coli* strain DH5 $\alpha$ , and recombinants were plated to provide approximately 800 colonies per plate and sufficient plates to provide approximately 50,000 total colonies per screen. Colonies were scraped from each plate, pooled, and plasmid DNA prepared from each pool. The pooled DNA was then used to transfect a sub-confluent layer of monkey COS-7 cells using DEAE-dextran followed by chloroquine treatment, as described by Luthman et al. (*Nucl. Acids Res.* 11:1295, 1983) and McCutchan et al. (*J. Natl. Cancer Inst.* 41:351, 1986). The cells were then grown in culture for three days to permit transient expression of the inserted sequences. After three days, cell culture supernatants were discarded and the cell monolayers in each plate assayed for

2133326

TNF binding as follows. Three ml of binding medium containing  $1.2 \times 10^{-11}$  M  $^{125}\text{I}$ -labeled FLAG<sup>®</sup>-TNF was added to each plate and the plates incubated at 4°C for 120 minutes. This medium was then discarded, and each plate was washed once with cold binding medium (containing no labeled TNF) and twice with cold PBS. The edges of each plate were then broken off, leaving a flat disk which was contacted with X-ray film for 72 hours at -70°C using an intensifying screen as described by Sims et al., *Science* 241:585 (1988). TNF binding activity was visualized on the exposed films as a dark focus against a relatively uniform background.

After approximately 240,000 recombinants from the library had been screened in this manner, one transfectant pool was observed to provide TNF binding foci which were clearly apparent against the background exposure. A frozen stock of bacteria from the positive pool was then used to obtain plates of approximately 150 colonies. Replicas of these plates were made on nitrocellulose filters, and the plates were then scraped and plasmid DNA prepared and transfected as described above to identify a positive plate. Bacteria from individual colonies from the nitrocellulose replica of this plate were grown in 0.2 ml cultures, which were used to obtain plasmid DNA, which was transfected into COS-7 cells as described above. In this manner, a single clone was isolated which was capable of inducing expression of human TNF-R in COS cells. The expression vector pCAV/NOT containing this TNF-R cDNA has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852, USA (Accession No. 68088) under the name pCAV/NOT-TNF-R.

#### Example 4

##### Construction of cDNAs Encoding Soluble huTNF-R $\Delta$ 235

25

A cDNA encoding a soluble huTNF-R $\Delta$ 235 was constructed. The encoded protein comprises the sequence of amino acids -22 to 235 of Figure 2A. Processing of the signal sequence yields a protein having the sequence of amino acids 1 to 235 of Figure 2A. An 840 bp fragment was excised from pCAV/NOT-TNF-R with the restriction enzymes Not1 and Pvu2. Not1 cuts at the multiple cloning site of pCAV/NOT-TNF-R and Pvu2 cuts within the TNF-R coding region 20 nucleotides 5' of the transmembrane region. In order to reconstruct the 3' end of the TNF-R sequences, two oligonucleotides were synthesized and annealed to create the following oligonucleotide linker:

35

SEQ ID NO: 9

Pvu2	BamH1 Bgl2
CTGAAGGGAGCACTGGCGACTAAGGATCCA	
GACTTCCCTCGTGACCGCTGATTCCCTAGGTCTAG	
AlaGluGlySerThrGlyAsp <u>End</u>	

5

This oligonucleotide linker has terminal Pvu2 and Bgl2 restriction sites, regenerates 20 nucleotides of the TNF-R, followed by a termination codon (underlined) and a BamH1 restriction site (for convenience in isolating the entire soluble TNF-R by Not1/BamH1 digestion). This oligonucleotide was then ligated with the 840 bp Not1/Pvu2 TNF-R insert into Bgl2/Not1 cut pCAV/NOT to yield psolhuTNF-R $\Delta$ 235/CAVNOT, which was transfected into COS-7 cells as described above. This expression vector induced expression of soluble human TNF-R which was capable of binding TNF.

15

### Example 5

#### Construction of cDNAs Encoding Soluble huTNF-R $\Delta$ 185

A cDNA encoding a soluble huTNF-R $\Delta$ 185 having the sequence of amino acids -22-185 of Figure 2A (or amino acids 1-185 upon processing of the signal sequence in an appropriate host cell) was constructed by excising a 640 bp fragment from pCAV/NOT-TNF-R with the restriction enzymes Not1 and Bgl2. Not1 cuts at the multiple cloning site of pCAV/NOT-TNF-R and Bgl2 cuts within the TNF-R coding region at nucleotide 637, which is 237 nucleotides 5' of the transmembrane region. The following oligonucleotide linkers were synthesized:

25

Bgl2	
SEQ ID NO: 10	5'-GATCTGTAACGTGGTGGCCATCCCTGGGAATGCAAGCATGGATGC-3'
	ACATTGCACCACCGGTAGGGACCCTTACGTTTCG
	IleCysAsnValValAlaIleProGlyAsnAlaSerMetAspAla

30

	Not1
SEQ ID NO: 11	5'-AGTCTGCACGTCCACGTCCCCCACC <u>GGTGAGC</u> -3'
	TACCTACGTCAGACGTGCAGGTGCAGGGGGTGGGCCACTCGCCGG
	ValCysThrSerThrSerProThrArgEnd

35

The above oligonucleotide linkers reconstruct the 3' end of the receptor molecule up to nucleotide 708, followed by a termination codon (underlined). These oligonucleotides were then ligated with the 640 bp Not1 TNF-R insert into Not1 cut pCAV/NOT to yield the expression vector psolTNFR $\Delta$ 185/CAVNOT, which was transfected into COS-7 cells as described above. This expression vector induced expression of soluble human TNF-R which was capable of binding TNF.

40

2133326

**Example 6**

Construction of cDNAs Encoding Soluble huTNF-RΔ163

5 A cDNA encoding a soluble huTNF-RΔ163 having the sequence of amino acids -22-163 of Figure 2A (1-163 upon processing of the signal sequence) was constructed by excising a 640 bp fragment from from pCAV/NOT-TNF-R with the restriction enzymes Not1 and Bgl2 as described in Example 4. The following oligonucleotide linkers were synthesized:

10

```

                Bgl2           Not1
SEQ ID NO: 12  5'-GATCTGTTGAGC   -3'
                ACAACTCGCCGG
                IleCysEnd
    
```

15

This above oligonucleotide linker reconstructs the 3' end of the receptor molecule up to nucleotide 642 (amino acid 163), followed by a termination codon (underlined). This oligonucleotide was then ligated with the 640 bp Not1 TNF-R insert into Not1 cut pCAV/NOT to yield the expression vector psolTNFRΔ163/CAVNOT, which was transfected into COS-7 cells as described above. This expression vector induced expression of soluble human TNF-R which was capable of binding TNF in the binding assay described in Example 1.

25

**Example 7**

Construction of cDNAs Encoding Soluble huTNF-RΔ142

30 A cDNA encoding a soluble huTNF-RΔ142 (having the sequence of amino acids -22-142 of Figure 2A (1-142 upon processing of the signal sequence) was constructed by excising a 550 bp fragment from from pCAV/NOT-TNF-R with the restriction enzymes Not1 and AlwN1. AlwN1 cuts within the TNF-R coding region at nucleotide 549. The following oligonucleotide linker was synthesized:

35

```

                Bgl2           Not1
SEQ ID NO: 13  5'-CTGAAACATCAGACGTGGTGTGCAAGCCCTGTTAAA-3'
                CTTGACTTTGTAGTCTGCACCACACGTTTCGGGACAATTTCTAGA
                End
    
```

This above oligonucleotide linker reconstructs the 3' end of the receptor molecule up to nucleotide 579 (amino acid 142), followed by a termination codon (underlined). This

oligonucleotide was then ligated with the 550 bp Not1/AlwN1 TNF-R insert into Not1/Bgl2 cut pCAV/NOT to yield the expression vector psolTNFR $\Delta$ 142/CAVNOT, which was transfected into COS-7 cells as described above. This expression vector did not induce expression of soluble human TNF-R which was capable of binding TNF. It is believed that this particular construct failed to express biologically active TNF-R because one or more essential cysteine residues (e.g., Cys<sup>157</sup> or Cys<sup>163</sup>) required for intramolecular bonding (for formation of the proper tertiary structure of the TNF-R molecule) was eliminated.

10

### Example 8

#### Isolation of Human type I IL-1R cDNA Clones

cDNA encoding a human type I IL-1R protein was isolated by hybridization to a probe derived from murine type I IL-1R cDNA. Cloning of this murine IL-1R cDNA is described in example 4 of EP 318,296. A vector containing the murine cDNA was deposited with the American Type Culture Collection under the name GEMBL78 on November 19, 1987 and given the accession number ATCC 67563.

A 2356 base pair (bp) fragment of this deposited murine clone 78 was isolated as described by Sims et al. (*Science* 241:585, 1988) and radiolabeled by nick-translation using DNA polymerase I for use as a probe. The method employed was substantially similar to that disclosed by Maniatis et al. (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982, p. 109).

The probe was used to screen human cDNA libraries for human IL-1R, as described by Sims et al., *Proc. Natl. Acad. Sci. (USA)* 86:8946, 1989. A cDNA library was constructed by reverse transcription of polyadenylated mRNA isolated from total RNA extracted from the cultured cells of a human T-cell line designated clone 22, described by Acres et al. (*J. Immunol.* 138:2132, 1987). These cells were cultured in RPMI 1640 medium plus 10% fetal bovine serum as described by Acres et al. (*supra*), in the presence of 10 ng/ml OKT3 antibody and 10 ng/ml human IL-2. The cDNA was rendered double-stranded using DNA polymerase I, blunt-ended with T4 DNA polymerase, methylated with EcoRI methylase to protect EcoRI cleavage sites within the cDNA, and ligated to EcoRI linkers. The resulting constructs were digested with EcoRI to remove all but one copy of the linkers at each end of the cDNA, and ligated to EcoRI-cut and dephosphorylated arms of bacteriophage  $\lambda$ gt10 (Huynh et al., *DNA Cloning: A Practical Approach*, Glover, ed., IRL Press, pp. 49-78). The ligated DNA



2133326

was packaged into phage particles using a commercially available kit (Stratagene Cloning Systems, San Diego, CA, USA 92121) to generate a library of recombinants. Recombinants were plated on *E. coli* strain C600(hf1-) and screened by standard plaque hybridization techniques under conditions of moderate stringency (50°C, 6 x  
5 SSC).

Following several rounds of screening, nine clones were isolated from the library which hybridized to the cDNA probe. The clones were plaque purified and used to prepare bacteriophage DNA which was digested with EcoRI. The digests were electrophoresed on an agarose gel, blotted onto nylon filters, and retested for  
10 hybridization. The clones were digested with EcoRI followed by preparative agarose gel electrophoresis, then subcloned into an EcoRI-cut derivative (pGEMBL) of the standard cloning vector pBR322 containing a polylinker having a unique EcoRI site, a BamHI site and numerous other unique restriction sites. An exemplary vector of this type is described by Dente et al. (*Nucl. Acids Res.* 11:1645, 1983).

15 Restriction mapping and sequencing of a 4.8 kb human IL-1R clone indicated that the clone included a sequence encoding 518 amino acids which exhibited 80% amino acid sequence identity to the corresponding murine sequence in the extracellular, or N-terminal region distal to the transmembrane region, 63% identity in the transmembrane region, and 87% identity in the cytoplasmic, or C-terminal region. A  
20 440 bp EcoRI-NsiI fragment derived from the 5' portion of the human IL-1R clone was <sup>32</sup>P-labeled by nick-translation as described above and used to screen a cDNA library produced by randomly-priming human T cell line clone 22 mRNA prepared as described above. 23 clones which hybridized to the probe were isolated and analyzed by restriction mapping. Sequencing of one of these clones provided the sequence  
25 information corresponding to the remaining N-terminal 34 amino acids of the human protein. The DNA and deduced amino acid sequence of the complete coding region of the type I human IL-1R are shown in SEQ ID NOS:5 and 6. This human IL-1R protein comprises 569 amino acids (including a 20 amino acid signal peptide), and includes 16 cysteine residues, 13 of which are conserved between the murine and human genes. In  
30 addition, the human sequence includes six potential N-glycosylation sites, of which five are conserved between murine and human.

**Example 9****Isolation of cDNAs Encoding Type II IL-1R**

5 A DNA sequence encoding human type II IL-1R was isolated from a cDNA library prepared using standard methods, by reverse transcription of polyadenylated RNA isolated from the human B cell lymphoblastoid line CB23, described by Benjamin & Dower, *Blood* 75:2017, 1990. Briefly, the CB23 cell line is an EBV-transformed cord blood (CB) lymphocyte cell line, which was derived using the methods described  
10 by Benjamin et al., *Proc. Natl. Acad. Sci. USA* 81:3547, 1984.

The CB23 library was screened by modified direct expression of pooled cDNA fragments in the monkey kidney cell line CV-1/EBNA-1 using a mammalian expression vector (pDC406) that contains origins of replication derived from SV40, Epstein-Barr virus and pBR322. pDC406 is a derivative of HAV-EO described by  
15 Dower et al., *J. Immunol.* 142:4314 (1989). pDC406 differs from HAV-EO by the deletion of the intron present in the adenovirus 2 tripartite leader sequence in HAV-EO. The CV-1/EBNA-1 cell line was derived by transfection of the CV-1 cell line with the gene encoding Epstein-Barr virus nuclear antigen-1 (EBNA-1) and with a vector containing CMV regulatory sequences, so that EBNA-1 is expressed under the control  
20 of the human CMV immediate-early enhancer/promoter. The EBNA-1 gene allows the episomal replication of expression vectors such as pDC406 that contain the EBV origin of replication.

Transfectants expressing biologically active type II IL-1R were initially identified using a modified slide autoradiographic technique, substantially as described  
25 by Gearing et al., *EMBO J.* 8:3667, 1989. Briefly, CV-1/EBNA-1 cells were transfected with miniprep DNA in pDC406 from pools of cDNA clones directly on glass slides and cultured for 2-3 days to permit transient expression of type II IL-1R. The slides containing the transfected cells were then incubated with medium containing <sup>125</sup>I-IL-1 $\beta$ , washed to remove unbound labeled IL-1 $\beta$ , fixed with gluteraldehyde, and  
30 dipped in liquid photographic emulsion and exposed in the dark. After developing the slides, they were individually examined with a microscope and positive cells expressing type II IL-1R were identified by the presence of autoradiographic silver grains against a light background.

Using this approach, approximately 250,000 cDNAs were screened in pools of  
35 approximately 3,000 cDNAs using the slide autoradiographic method until assay of one transfectant pool showed multiple cells clearly positive for IL-1 $\beta$  binding. This pool

2133326

was then partitioned into pools of 500 and again screened by slide autoradiography. A positive pool was identified. This pool was further partitioned into pools of 75 and screened by plate binding assays analyzed by quantitation of bound  $^{125}\text{I}$ -IL-1 $\beta$ . The cells were scraped off and counted to determine which pool of 75 was positive.

5 Individual colonies from this pool of 75 were screened until a single clone was identified which directed synthesis of a surface protein with detectable IL-1 $\beta$  binding activity. This clone was isolated, and its insert was sequenced to determine the sequence of the human type II IL-1R cDNA that is presented along with the amino acid sequence encoded thereby in SEQ ID NOS:7 and 8. The pDC406 cloning vector

10 containing the human type II IL-1R cDNA, designated pHu IL-1R-II 75, was deposited in *E.coli* host cells with the American Type Culture Collection, Rockville, MD. USA (ATCC) on June 5, 1990 under accession number ATCC 68337. The deposit was made under the conditions of the Budapest Treaty.

Like most mammalian genes, mammalian type II IL-1R is presumably encoded

15 by multi-exon genes. Alternative mRNA constructs which can be attributed to different mRNA splicing events following transcription, and which share large regions of identity or similarity with the cDNAs claimed herein, are considered to be within the scope of the present invention.

20

### Example 10

#### Construction and Expression of cDNAs Encoding Human Soluble Type II IL-1R

A cDNA encoding a soluble human type II IL-1R (having the sequence of amino acids

25 -13-333 of SEQ ID NO:8) was constructed by polymerase chain reaction (PCR) amplification using the full length type II IL-1R cDNA clone 75 (ATCC 68337) in vector pDC406 (described in example 9) as a template. The following 5' oligonucleotide primer (SEQ ID NO:14) and 3' oligonucleotide primer (SEQ ID NO:15) were first constructed:

30

SEQ ID NO: 14 5'-GCGTCGACCTAGTGACGCTCATACAAATC-3'  
<SalI>

35

SEQ ID NO: 15 5'-GCGCGGCCGCTCAGGAGGAGGCTTCCTTGACTG-3'  
<-NotI->End\1191 \1172

The 5' primer corresponds to nucleotides 31-51 from the untranslated region of human type II IL-1R clone 75 (SEQ ID NO:7) with a 5' add-on of a Sall restriction site; this nucleotide sequence is capable of annealing to the (-) strand complementary to nucleotides 31-51 of human clone 75. The 3' primer is complementary to nucleotides 1191-1172 (which includes anti-sense nucleotides encoding 3 amino acids of human type II IL-1R clone 75 and has a 5' add-on of a NotI restriction site and a stop codon.

The following PCR reagents were added to a 1.5 ml Eppendorf microfuge tube: 10  $\mu$ l of 10X PCR buffer (500 mM KCl, 100 mM Tris-HCl, pH 8.3 at 25°C, 15 mM MgCl<sub>2</sub>, and 1 mg/ml gelatin) (Perkins-Elmer Cetus, Norwalk, CN), 10  $\mu$ l of a 2mM solution containing each dNTP (2 mM dATP, 2 mM dCTP, 2 mM dGTP and 2mM dTTP), 2.5 units (0.5  $\mu$ l of standard 5000 units/ml solution) of *Taq* DNA polymerase (Perkins-Elmer Cetus), 50 ng of template DNA and 5  $\mu$ l of a 20  $\mu$ M solution of each of the above oligonucleotide primers and 74.5  $\mu$ l water to a final volume of 100  $\mu$ l. The final mixture was then overlaid with 100  $\mu$ l parafin oil. PCR was carried out using a DNA thermal cycler (Ericomp, San Diego, CA) by initially denaturing the template at 94° for 90 seconds, reannealing at 55° for 75 seconds and extending the cDNA at 72° for 150 seconds. PCR was carried out for an additional 20 cycles of amplification using a step program (denaturation at 94°, 25 sec; annealing at 55°, 45 sec; extension at 72°, 150 sec.), followed by a 5 minute extension at 72°.

The sample was removed from the parafin oil and DNA extracted by phenolchloroform extraction and spun column chromatography over G-50 (Boehringer Mannheim). A 10  $\mu$ l aliquot of the extracted DNA was separated by electrophoresis on 1% SeaKem agarose (FMC BioProducts, Rockland, ME) and stained with ethidium bromide to confirm that the DNA fragment size was consistent with the predicted product.

20  $\mu$ l of the PCR-amplified cDNA products were then digested with Sall and NotI restriction enzymes using standard procedures. The Sall/NotI restriction fragment was then separated on a 1.2% Seaplaque™ low gelling temperature (LGT) agarose, and the band representing the fragment was isolated. The fragment was ligated into the pDC406 vector by a standard "in gel" ligation method. The resulting vector was transfected into CV1-EBNA cells and the soluble IL-1R protein was expressed.

**Example 11****Construction of Vector Encoding Di-TNF-R**

5 A vector encoding a fusion protein of the formula TNF-R - peptide linker - TNF-R and depicted in figure 3 was constructed as follows. Pertinent restriction enzyme cleavage sites in the TNF-R sequence are also shown in Figure 2.

The expression vector constructed in example 4 and designated psol huTNF-  
10 RΔ235/CAVNOT was digested with the restriction enzyme Not I, which cuts at the multiple cloning site of the pCAV/NOT vector (i.e., upstream of the TNF-R sequence inserted therein). The overhang generated by Not I digestion was filled in using the Klenow fragment of DNA polymerase I to produce a blunt end. The vector was then digested with Bam HI which cleaves downstream of the stop codon that follows the  
15 codon for amino acid 235, as shown in example 4.

The blunted NotI/Bam HI fragment containing the TNF-R sequence was isolated by conventional procedures and inserted into a plasmid vector designated pCAV/DHFR which had been digested with Sma I and Bgl II. The pCAV/DHFR vector is an expression vector containing SV40 promoter sequences upstream of a  
20 multiple cloning site and other features as described for pCAV/NOT in example 3, and also contains a dihydrofolate reductase (DHFR) gene as a selectable marker. The DHFR gene confers a selective advantage on otherwise DHFR<sup>-</sup> mammalian cells that have taken up the vector, when grown in the presence of methotrexate (MTX). Sma I digestion produces blunt ends, to which the blunted Not I ends of the TNF-R-containing fragment are ligated. The Bgl II-generated overhangs are ligated to the Bam  
25 HI-digested ends of the TNF-R-containing fragment. The ligation destroys the Bam HI and Bgl II sites. *E. coli* cells are transformed with the ligation mixture by conventional procedures. Plasmid DNA is recovered from the host cells and the desired construct is confirmed by restriction analysis. The resulting vector containing the TNF-R insert is  
30 designated pCAV/DHFR/TNF-R.

The following DNA fragments were isolated for use in preparing a vector encoding two TNF-R polypeptides separated by a peptide linker:

(A) Asp718 (a restriction enzyme) to Esp I fragment of the pCAV/DHFR/TNF-R vector. Asp 718 cleaves the vector upstream of the inserted TNF-R sequence; EspI  
35 (available from U.S. Biochemicals) cleaves the TNF-R sequence at the position shown in Figure 2. The desired fragment is about 6.7 kilobase-pairs (kbp) in length and includes vector sequences and the 3' end of a TNF-R sequence.

(B) Asp718 to PvuII fragment of the expression vector constructed in example 4 and designated psol hu TNF-R  $\Delta$ 235/CAVNOT. Asp718 cleaves the vector upstream of the inserted TNF-R sequence. The desired 865 bp fragment includes a TNF-R sequence extending from the 5' signal sequence through the Pvu II site shown in example 4.

(C) a double-stranded oligonucleotide having the sequence (SEQ ID NOS:16 and 17):

5' CTGAAGGGAGCACTGGCGACGGTGGCGGTGGATCCGGCGGTGGCGGCGGCTCATTGCCCGCCCAGG 3'

3' GACTTCCCTCGTGACCGCTGCCACCGCCACCTAGGCCGCCACCGCCGCCGAGTAACGGGCGGG 5'

GluGlySerThrGlyAspGlyGlyGlyGlySerGlyGlyGlyGlyGlySerLeuProAlaGlnVal

(D) Bgl I to Esp I fragment of the expression vector constructed in example 4 and designated psol hu TNF-R  $\Delta$ 235/CAVNOT. The desired fragment is about 304 bp in length. Bgl I cleaves within the codon for amino acid 5 (Val) of TNF-R; Esp I cleaves the TNF-R sequence at the position shown in Figure 2; the remainder (3' end) of this soluble TNF-R-encoding sequence is provided by fragment (A).

The oligonucleotide (C) is prepared by conventional procedures for chemical synthesis of oligonucleotides. The oligonucleotide reconstructs the 3' end of the first TNF-R sequence from the Pvu II site through the last amino acid of the extracellular domain (Asp at position 235). The oligonucleotide also contains an in-frame sequence encoding the peptide linker Gly<sub>4</sub>SerGly<sub>5</sub>Ser. The sequence of this portion of the oligonucleotide may be varied if desired to encode other peptide linkers. Oligonucleotide C also reconstructs the 5' end of the second TNF-R sequence from a codon for leucine, the first amino acid of the mature protein, through a partial codon for valine (amino acid 5) within the protruding 3' overhang, which will regenerate the Val codon when ligated to the complementary overhang on the Bgl I-digested end of fragment D.

The DNA fragments designated A-D were ligated together in the positions shown in figure 3 to form a vector designated pCAV DHFR Di-TNF-R which encodes a fusion protein of the present invention. *E. coli* cells are transformed with the ligation mixture by conventional procedures. Plasmid DNA is recovered from the host cells and the desired construct is confirmed by restriction analysis. The upstream TNF-R polypeptide encoded by this expression vector contains amino acids -22 to 235 of SEQ ID NO:2 (i.e., a soluble TNF-R including the N-terminal signal sequence and the entire extracellular domain without a stop codon). The downstream TNF-R polypeptide lacks

2133326

42

the signal sequence and contains amino acids 1-235 of SEQ ID NO:2, with a stop codon positioned immediately after amino acid 235. The peptide linker of this particular construct is Gly<sub>4</sub>SerGly<sub>5</sub>Ser.

5 Mammalian cells are transfected with the expression vector by conventional procedures and cultured to produce the desired fusion protein. One suitable mammalian cell line is a DHFR<sup>-</sup> Chinese hamster ovary cell line designated CHO-K1 and available from the American Type Culture Collection, Rockville, MD, under accession number CCL61. The cells may be transfected with the expression vector by standard calcium phosphate precipitation, essentially as described by Graham and van der Eb, *Virology* 10 52:456 (1983). The transfected cells are cultured under suitable conventional conditions, and the presence of the desired fusion protein in the culture medium is confirmed by assays such as those described in examples 1 and 2.

Another suitable mammalian cell line is the COS-7 cell line. In one experiment, COS-7 cells were transfected with the TNF-R dimer-encoding expression vector 15 described above and cultured to allow expression of the dimer. Unconcentrated supernatant (containing the secreted dimer) was assayed for the ability to inhibit binding of radioiodinated murine TNF $\alpha$  to U937 cell (a human monocyte-like cell line bearing endogenous receptors for TNF). The inhibition binding assay was conducted according to conventional procedures, similar to those described in example 2, section 20 C. Negative control tubes (to measure non-specific binding) contained non-radiolabeled TNF $\alpha$ , U937 cells, and radioiodinated TNF $\alpha$ . Positive control tubes (to measure maximum, uninhibited binding of radioiodinated TNF $\alpha$  to the U937 cells) contained binding medium, U937 cells, and radioiodinated TNF $\alpha$ . The dimer-containing supernatant inhibited over 90% of the TNF binding to U937 cells that was 25 seen for the positive control.

### Example 12

#### Fusion Protein Comprising One IL-1R Polypeptide and Two TNF-R Polypeptides

30

A plasmid vector containing DNA encoding a fusion protein of the formula IL-1R-peptide linker-TNF-R-peptide linker-TNF-R is constructed as follows.

cDNA encoding a soluble type I IL-1R polypeptide was isolated and amplified using the well known polymerase chain reaction (PCR) procedure. The following 35 oligonucleotides were synthesized for use as primers in the PCR reaction:

SEQ ID NO: 18 5' ACCGAGGGACCTGAGCG 3'

SEQ ID NO: 19 3' TCAATTATATAGGTCAGTGACCACCGCCACCTAGGCCGCCACCGC  
CGCCGAGT 5'

5 These oligonucleotides as well as those discussed below are synthesized by conventional procedures, e.g., by using an automated DNA synthesis machine such as those available from Biosearch, Inc., San Rafael, California or Applied Biosystems. The template employed in the PCR reaction is a plasmid vector prepared by inserting type I human IL-1R cDNA into a vector designated SF CAV. The SF CAV vector is a mammalian expression vector shown in figure 5 (which depicts the use of SF CAV in an additional vector construction described below.) SF CAV in *E. coli* cells was  
10 deposited with the American Type Culture Collection on February 27, 1992, under the terms of the Budapest Treaty, and was given accession number 68922.

The SV40, CMV, pA, and VA sequences and ampicillin resistance gene in SF CAV are as described for pCAV/NOT in example 3 above and also in example 8 and figure 3 of PCT application WO 90/05183. A multiple cloning site positioned between  
15 the adenovirus-2 tripartite leader (TPL) and pA sequences contains recognition sites for the restriction endonucleases XhoI, KpnI, SmaI, NotI, and BglII. The TPL sequence differs from that of pCAV/NOT in that a region believed to be detrimental to construction of IL-1R-encoding vectors has been deleted from the SF CAV TPL sequence. The adverse impact on IL-1R vectors may possibly be attributable to a  
20 cryptic promoter in the undesirable sequence, from which undesired protein is expressed in *E. coli*. Low level expression of human IL-1R off the cryptic promoter may be toxic to the bacteria.

SF CAV is digested with SmaI, which recognizes a unique restriction site within the multiple cloning site and produces blunt ends. A DNA fragment containing  
25 type I IL-1R cDNA is produced by StyI/BglII digestion followed by filling in the overhangs using the Klenow fragment of DNA polymerase I to generate blunt ends. StyI cleaves at nucleotide 49 and BglII cleaves at nucleotide 1997 of SEQ ID NO:5. The IL-1R cDNA fragment is ligated into the SmaI-digested SF CAV vector, and *E. coli* cells are transformed with the ligation mixture by standard procedures. The  
30 resulting vector is recovered from the *E. coli* cells and used as the template in the PCR reaction.

The 5' primer (SEQ ID NO:18) corresponds to a 17-nucleotide sequence found in the vector upstream of the inserted IL-1R cDNA. The 3' primer (SEQ ID NO:19) includes a segment complementary to nucleotides 1060-1079 of SEQ ID NO:5, which  
35 encode amino acids 306 (partial codon) through 312, near the C-terminus of the IL-1R



2133326

extracellular domain. This 3' primer also contains a sequence encoding the peptide linker Gly<sub>4</sub>SerGly<sub>5</sub>Ser.

A PCR reaction is conducted using any suitable procedure, such as those described in Sarki et al., *Science* 239:487 (1988); in *Recombinant DNA Methodology*, Wu et al., eds., Academic Press Inc., San Diego (1989), pp. 189-196; and in *PCR Protocols: A Guide to Methods and Applications*, Innis et al., eds., Academic Press, Inc. (1990). An example of a suitable PCR procedure is as follows. All temperatures are in degrees centigrade. The following PCR reagents are added to a 0.5 ml Eppendorf microfuge tube: 10 µl of 10X PCR buffer (500 mM KCl, 100 mM Tris-HCl, pH 8.3 at 25°C, 25 mM MgCl<sub>2</sub>, and 1 mg/ml gelatin) (Perkin-Elmer Cetus, Norwalk, CN), 8 µl of a 2.5 mM solution containing each dNTP (2 mM dATP, 2 mM dCTP, 2 mM dGTP and 2 mM dTTP), 2.5 units (0.5 µl of standard 5000 units/ml solution) of *Taq* DNA polymerase (Perkins-Elmer Cetus), 1 ng of template DNA, 100 picomoles of each of the oligonucleotide primers, and water to a final volume of 100 µl. The final mixture is then overlaid with 100 µl parafin oil. PCR is carried out using a DNA thermal cycler (Ericomp, San Diego, CA). The template is denatured at 94° for 5 minutes and PCR is carried out for 25 cycles of amplification using a step program (denaturation at 94°, 1.5 minutes; annealing at 60°, 1 minute; extension at 72°, 1 minute).

Electrophoresis of an aliquot of the reaction mixture on 1% SeaKem low melting temperature (LMT) agarose (FMC BioProducts, Rockland, ME) and staining with ethidium bromide visualizes a single DNA fragment of the expected size as the PCR reaction product. The PCR-amplified DNA fragment comprises a short vector sequence that includes an Asp718 restriction site, upstream of a sequence encoding IL-1R amino acids 1(Asp) to 312 (Thr), followed by a single stranded segment encoding the peptide linker.

A second PCR reaction is conducted to isolate and amplify a cDNA fragment encoding a soluble TNF-R polypeptide. The template was the vector designated psolhuTNF-R Δ235/CAVNOT in example 4, which contains a cDNA insert encoding TNF-R amino acids -22 to 235 of SEQ ID NO:1. The primers employed in the reaction are:

SEQ ID NO: 20 5' GGTGGCGGTGGATCCGGCGGTGGCGGCGGCTCATTGCCCGCCAGGTGGCA 3'

35 SEQ ID NO: 21 3' TGACGCGCGACTCGTTTCG 5'

The 5' primer (SEQ ID NO:20) comprises a peptide linker encoding segment complementary to the peptide linker encoding portion of the SEQ ID NO:19 primer. The linker-encoding segment is followed by codons for the first six amino acids of mature TNF-R (Leu through Ala).

5 The 3' primer (SEQ ID NO:21) comprises nucleotides complementary to nucleotides 461-478 of TNF-R SEQ ID NO:1, which encode amino acids 103 (Tyr, partial codon) through 109 (Gln, partial codon). This primer also encompasses an *EspI* restriction site that is naturally present in this portion of the TNF-R protein.

10 The PCR reaction procedure is as described above. The DNA fragment amplified by this second PCR reaction comprises the above-described linker-encoding segment followed by a sequence encoding amino acids 1 (Leu) through 109 (Gln, partial codon) of TNF-R. This DNA fragment is visualized by electrophoresis followed by ethidium bromide staining of the gel, as described above. A third PCR reaction is conducted to isolate an amplified double stranded DNA fragment comprising  
15 IL-1R and TNF-R sequences separated by the linker sequence. The following reagents were combined in a 0.5 ml. tube:

20 3  $\mu$ l (about 5-10 ng) of the IL-1R-peptide linker DNA fragment amplified in the first PCR reaction above - a 3  $\mu$ l aliquot is taken directly from the LMT agarose by micropipette, using UV light to visualize the desired band on the ethidium bromide stained gel

25 3  $\mu$ l (about 5-10 ng) of the peptide linker-TNF-R DNA fragment amplified in the second PCR reaction above - a 3  $\mu$ l aliquot is micropipetted directly from the region of the LMT agarose gel that contains the desired band

30 10  $\mu$ l of 10 X PCR buffer (described above)  
100 pmole of the SEQ ID NO:18 oligonucleotide as the 5' primer  
100 pmole of the SEQ ID NO:21 oligonucleotide as the 3' primer  
4  $\mu$ l of a 2.5 mM solution containing each of the four dNTPs  
0.5  $\mu$ l of *Taq* DNA polymerase (5 units/ $\mu$ l)  
water to a final volume of 100  $\mu$ l

35 The PCR reaction cycles are conducted at the temperatures and for the time periods specified above. After the initial denaturing step, the complementary peptide linker-encoding segments of the two DNA fragments anneal. The end product of the reaction is a blunt-ended double-stranded amplified DNA fragment about 1370 bp in

length, comprising an IL-1R DNA sequence upstream of a sequence encoding a peptide linker, followed by a TNF-R DNA sequence.

A 25  $\mu$ l aliquot of this third PCR reaction mixture is, without purification, reacted with the restriction enzymes Asp718 and EspI. Asp 718 cleaves upstream of the IL-1R sequence, as discussed above. EspI cleaves within the TNF-R sequence, as shown in Figure 2 and discussed above. The restriction endonuclease Cell II is an isoschizomer of EspI and may be substituted for EspI. The desired fragment, referred to as fragment E hereinafter, is purified by conventional procedures such as separation by gel electrophoresis, e.g., on a 1.0% Seaplaque\* low melting temperature agarose gel, and isolation of the band representing the desired fragment.

Two additional DNA fragments designated F and G are isolated and joined with fragment E to construct an expression vector having a second TNF-R sequence fused (via a peptide linker sequence) downstream of the IL-1R-linker-TNF-R DNA fragment prepared above. The resulting vector and the positions of fragments E, F, and G contained therein are depicted in figure 4.

The DNA fragment designated F is prepared by digesting the vector depicted in Figure 3 and constructed in example 11 with EspI. A fragment containing a 3' portion of TNF-R (extending from the EspI site shown in Figure 2 through a codon for amino acid 235) followed by a sequence encoding a Gly<sub>4</sub>SerGly<sub>5</sub>Ser peptide linker that is followed by a second TNF-R sequence extending from the codon for amino acid 1 (Leu) to the EspI site in the second (downstream) TNF-R sequence of the figure 3 vector, is isolated. This fragment, about 739 base pairs in length, is designated fragment F hereinafter.

Fragment G, containing vector sequences (including DHFR) and a 3' portion of a TNF-R sequence is isolated from a "splice free" vector as follows. The pCAV/DHFR/TNF-R vector prepared in example 11 contains a sequence which is believed to be disadvantageous for construction of IL-1R-encoding vectors, possibly because of a cryptic promoter within this sequence from which undesired protein is expressed in *E. coli*. A derivative of pCAV/DHFR/TNF-R was prepared by replacing an NdeI/Asp 718 vector fragment that contains the undesirable sequence with an NdeI/Asp 718 fragment from splice free vector SF CAV (ATCC 68922, described above). The construction is depicted in figure 5.

SF CAV is digested with NdeI and Asp 718 (unique sites in this plasmid) and the fragment of about 500 bp labeled H in figure 5 is isolated. Fragment H lacks the undesired sequence found in the corresponding fragment of pCAV/DHFR/TNF-R.

\*Trade-mark

The pCAV/DHFR/TNF-R vector prepared in example 11 and shown in figure 5 is digested with Asp 718, EspI, and NdeI. An Asp 718/EspI fragment of about 495 bp is isolated (fragment I). An NdeI/EspI fragment of about 4647 bp is also isolated (fragment J).

5 DNA fragments H, I, and J prepared above are ligated together to form the splice free vector SF CAV/DHFR/TNF-R shown in figure 5. *E. coli* cells are transformed with the ligation mixture by conventional procedures. Plasmid DNA is recovered from the host cells and the desired construct is confirmed by restriction analysis. SF pCAV/DHFR/TNF-R is then digested with Asp718 and EspI. The  
10 fragment designated G in figure 5 contains vector sequences (including DHFR) and the 3' end of a TNF-R sequence (from the internal EspI site through the codon for amino acid 235 followed by a stop codon) and is isolated by conventional procedures.

DNA fragments E, F, and G prepared above are ligated together to form the vector depicted in figure 4 (and designated SF CAV DHFR tri-R). *E. coli* cells are  
15 transformed with the ligation mixture and the desired plasmid is recovered as described above. The fusion protein encoded by this vector comprises (from N- to C- terminus) amino acids -20 to 312 of type I IL-1R (SEQ ID NO:5); a Gly<sub>4</sub>SerGly<sub>5</sub>Ser peptide linker; amino acids 1 to 235 of TNF-R (SEQ ID NO:1); a Gly<sub>4</sub>SerGly<sub>5</sub>Ser peptide linker; and a second TNF-R polypeptide comprising amino acids 1 to 235 of SEQ ID  
20 NO:1.

Mammalian cells are transfected with the expression vector by conventional procedures and cultured to produce the desired fusion protein. One suitable mammalian cell line is a DHFR<sup>-</sup> Chinese hamster ovary cell line designated CHO-K1 and available from the American Type Culture Collection, Rockville, MD, under accession number  
25 CCL61. The cells may be transfected with the expression vector by standard calcium phosphate precipitation, essentially as described by Graham and van der Eb, *Virology* 52:456 (1983). The transfected cells are cultured under suitable conventional conditions, and the presence of the desired fusion protein in the culture medium is confirmed by assays such as those described in examples 1 and 2.

30 DNA sequencing of an expression vector constructed as described above revealed two point mutations in the fusion protein-encoding sequences, which may have occurred during PCR. The "T" at position 437 of SEQ ID NO:5 (type I IL-1R) was changed to a "C" and the "C" at position 687 of SEQ ID NO:1 (TNF-R) was changed to a "T". These mutations are silent, so the amino acid sequence of the IL-1R-peptide linker-TNF-R-peptide linker-TNF-R fusion protein (termed the "trimeric  
35 receptor" hereinafter) is as described above.

COS-7 cells were transfected with the expression vector and cultured to allow expression of the trimeric receptor. 5X concentrated supernatant (containing the secreted trimeric receptor protein) was assayed for the ability to inhibit binding of radioiodinated murine TNF $\alpha$  (0.5 nM) to U937 cells (a human monocyte-like cell line bearing endogenous receptors for TNF). The inhibition binding assay was conducted according to conventional procedures, similar to those described in example 2, section C. Negative control tubes (to measure non-specific binding) contained non-radiolabeled TNF $\alpha$ , U937 cells, and radioiodinated TNF $\alpha$ . Positive control tubes (to measure maximum, uninhibited binding of radioiodinated TNF $\alpha$  to the U937 cells) contained binding medium, U937 cells, and radioiodinated TNF $\alpha$ . The trimeric receptor-containing supernatant inhibited about 100% of the TNF binding to U937 cells that was seen for the positive control.

The 5X concentrated trimeric receptor-containing supernatant also was assayed for the ability to inhibit binding of radioiodinated human IL-1 $\alpha$  (0.14 nM) to EL4 6.1 cells, which bear endogenous IL-1 receptors. The EL4 6.1 cell line was derived from a murine thymoma cell line as described by MacDonald et al. (*J. Immunol.* 135:3944, 1985). Negative control tubes (to measure non-specific binding) contained non-radiolabeled IL-1 $\alpha$ , EL4 6.1 cells, and radioiodinated IL-1 $\alpha$ . Positive control tubes (to measure maximum, uninhibited binding of radioiodinated IL-1 $\alpha$  to the EL4 6.1 cells) contained binding medium, EL4 6.1 cells, and radioiodinated IL-1 $\alpha$ . The trimeric receptor-containing supernatant inhibited 47% of the IL-1 binding to EL4 6.1 cells that was seen for the positive control. Supernatant from COS-7 cells expressing a monomeric soluble type I IL-1R, assayed as a control, inhibited 52% of the IL-1 binding to EL4 6.1 cells that was seen for the positive control.

### Example 13

#### Fusion Protein Comprising One IL-1R Polypeptide and One TNF-R Polypeptide

Fragments E and G prepared in example 12 may be ligated together to form a vector containing a DNA sequence that encodes a fusion protein of the formula IL-1R-peptide linker-TNF-R, wherein the peptide linker is Gly<sub>4</sub>SerGly<sub>5</sub>Ser. IL-1R and TNF-R are soluble polypeptides as described in example 12. The fusion protein of example 12 is preferred due to the enhanced TNF binding achieved when two, rather than one, TNF-R polypeptides are employed.

COS-7 cells were transfected with the expression vector and cultured to allow expression of the IL-1R-peptide linker-TNF-R fusion protein. 5X concentrated supernatant (containing the secreted fusion protein) was assayed for the ability to inhibit binding of radioiodinated human IL-1 $\alpha$  (0.14 nM) to EL4 6.1 cells (a murine thymoma-derived cell line bearing IL-1 receptors as described in example 12). The inhibition binding assay procedure was similar to that described in example 2, section C, and the positive and negative controls were as described in example 12. The fusion protein-containing supernatant inhibited 53% of the IL-1 binding to EL4 6.1 cells that was seen for the positive control.

10

#### Additional Fusion Proteins

The skilled artisan will appreciate that the techniques disclosed herein may be employed to produce additional fusion proteins of the present invention, beyond those illustrative embodiments presented in the foregoing examples. The peptide linker may be varied by synthesizing an oligonucleotide encoding a different peptide linker sequence, for example. Further, alternative fragments of the IL-1R or TNF-R proteins may be isolated by employing PCR primers that anneal to a different desired portion of the disclosed DNA sequences thus defining the termini of the desired fragment. Oligonucleotides used to regenerate a terminus of a DNA fragment (e.g., the oligonucleotide employed in example 4) may be varied to position a stop codon after any desired amino acid. Further, the choice of expression vector will depend upon the intended host cells.

25

#### BRIEF DESCRIPTION OF THE SEQUENCE LISTING

SEQ ID NO:1 and SEQ ID NO:2 show the nucleotide sequence and encoded amino acid sequence of a human TNF-R cDNA. The mature protein is defined by amino acids 1 - 439. The signal peptide is defined by amino acids -22 through -1. The transmembrane region is defined by amino acids 236 - 265.

30

SEQ ID NO:3 and SEQ ID NO:4 show the nucleotide sequence and encoded amino acid sequence of the coding region of a human TNF-R cDNA. This TNF-R is a different protein from the TNF-R of SEQ ID NOS:1 and 2. The mature protein is defined by amino acids 1 - 415. The signal peptide is defined by amino acids -40 through -1. The transmembrane region is defined by amino acids 172 - 192.

35

SEQ ID NO:5 and SEQ ID NO:6 show the nucleotide sequence and encoded amino acid sequence of a human type I IL-1R cDNA. The mature protein is defined by amino acids 1 - 549. The predicted signal peptide is defined by amino acids -20 through -1. The transmembrane region is defined by amino acids 317 - 336.

5 SEQ ID NO:7 and SEQ ID NO:8 show the nucleotide sequence and encoded amino acid sequence of a human type II IL-1R cDNA. The mature protein is defined by amino acids 1 - 385. The predicted signal peptide is defined by amino acids -13 through -1. The transmembrane region is defined by amino acids 331 - 356.

10 SEQ ID NOS:9-15 and 18-21 depict oligonucleotides employed in constructing various recombinant plasmids, as described in the examples section.

SEQ ID NOS:16 and 17 depict an oligonucleotide and the amino acid sequence encoded thereby, which includes a Gly<sub>4</sub>SerGly<sub>5</sub>Ser peptide linker sequence. This oligonucleotide is employed in the vector construction described in example 11.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Smith, Craig A.
- (ii) TITLE OF INVENTION: Fusion Proteins Comprising Tumor Necrosis Factor Receptor
- (iii) NUMBER OF SEQUENCES: 21
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Immunex Corporation
  - (B) STREET: 51 University Street
  - (C) CITY: Seattle
  - (D) STATE: Washington
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 98101
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 479,661
  - (B) FILING DATE: 07-FEB-1990
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 523,635
  - (B) FILING DATE: 10-MAY-1990
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 701,415
  - (B) FILING DATE: 16-JUN-1991
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 821,716
  - (B) FILING DATE: 14-JAN-1992
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Seese, Kathryn A.
  - (B) REGISTRATION NUMBER: 32,172
  - (C) REFERENCE/DOCKET NUMBER: 2502
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (206) 587-0430
  - (B) TELEFAX: (206) 587-0606

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1641 base pairs
  - (B) TYPE: nucleic acid



2133326

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: Homo sapiens  
(G) CELL TYPE: Fibroblast  
(H) CELL LINE: WI-26 VA4
- (vii) IMMEDIATE SOURCE:  
(A) LIBRARY: WI-26 VA4  
(B) CLONE: 1
- (ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 88..1473
- (ix) FEATURE:  
(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 154..1470
- (ix) FEATURE:  
(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 88..153
- (x) PUBLICATION INFORMATION:  
(A) AUTHORS: Smith, Craig A.  
Davis, Terri  
Anderson, Dirk  
Solam, Lisabeth  
Beckmann, M. P.  
Jerzy, Rita  
Dower, Steven K.  
Cosman, David  
Goodwin, Raymond G.  
(B) TITLE: A Receptor for Tumor Necrosis Factor Defines  
an Unusual Family of Cellular and Viral Proteins  
(C) JOURNAL: Science  
(D) VOLUME: 248  
(E) PAGES: 1019-1023  
(G) DATE: 25-MAY-1990
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCGAGGCAGG CAGCCTGGAG AGAAGGCGCT GGGCTGCGAG GGCGCGAGGG CGCSAGGGCA 60

GGGGGCAACC GGACCCCGCC CGCATCC ATG GCG CCC GTC GCC GTC TGG GCC 111  
Met Ala Pro Val Ala Val Trp Ala  
-22 -20 -15

GCG CTG GCC GTC GGA CTG GAG CTC TGG GCT GCG GCG CAC GCC TTG CCC 159  
Ala Leu Ala Val Gly Leu Glu Leu Trp Ala Ala Ala His Ala Leu Pro  
-10 -5

SUBSTITUTE SHEET

GCC CAG GTG GCA TTT ACA CCC TAC GCC CCG GAG CCC GGG AGC ACA TGC Ala Gln Val Ala Phe Thr Pro Tyr Ala Pro Glu Pro Gly Ser Thr Cys 5 10 15	207
CGG CTC AGA GAA TAC TAT GAC CAG ACA GCT CAG ATG TGC TGC AGC AAA Arg Leu Arg Glu Tyr Tyr Asp Gln Thr Ala Gln Met Cys Cys Ser Lys 20 25 30	255
TGC TCG CCG GGC CAA CAT GCA AAA GTC TTC TGT ACC AAG ACC TCG GAC Cys Ser Pro Gly Gln His Ala Lys Val Phe Cys Thr Lys Thr Ser Asp 35 40 45 50	303
ACC GTG TGT GAC TCC TGT GAG GAC AGC ACA TAC ACC CAG CTC TGG AAC Thr Val Cys Asp Ser Cys Glu Asp Ser Thr Tyr Thr Gln Leu Trp Asn 55 60 65	351
TGG GTT CCC GAG TGC TTG AGC TGT GGC TCC CGC TGT AGC TCT GAC CAG Trp Val Pro Glu Cys Leu Ser Cys Gly Ser Arg Cys Ser Ser Asp Gln 70 75 80	399
GTG GAA ACT CAA GCC TGC ACT CGG GAA CAG AAC CGC ATC TGC ACC TGC Val Glu Thr Gln Ala Cys Thr Arg Glu Gln Asn Arg Ile Cys Thr Cys 85 90 95	447
AGG CCC GGC TGG TAC TGC GCG CTG AGC AAG CAG GAG GGG TGC CGG CTG Arg Pro Gly Trp Tyr Cys Ala Leu Ser Lys Gln Glu Gly Cys Arg Leu 100 105 110	495
TGC GCG CCG CTG CGC AAG TGC CGC CCG GGC TTC GGC GTG GCC AGA CCA Cys Ala Pro Leu Arg Lys Cys Arg Pro Gly Phe Gly Val Ala Arg Pro 115 120 125 130	543
GGA ACT GAA ACA TCA GAC GTG GTG TGC AAG CCC TGT GCC CCG GGG ACG Gly Thr Glu Thr Ser Asp Val Val Cys Lys Pro Cys Ala Pro Gly Thr 135 140 145	591
TTC TCC AAC ACG ACT TCA TCC ACG GAT ATT TGC AGG CCC CAC CAG ATC Phe Ser Asn Thr Thr Ser Ser Thr Asp Ile Cys Arg Pro His Gln Ile 150 155 160	639
TGT AAC GTG GTG GCC ATC CCT GGG AAT GCA AGC ATG GAT GCA GTC TGC Cys Asn Val Val Ala Ile Pro Gly Asn Ala Ser Met Asp Ala Val Cys 165 170 175	687
ACG TCC ACG TCC CCC ACC CGG AGT ATG GCC CCA GGG GCA GTA CAC TTA Thr Ser Thr Ser Pro Thr Arg Ser Met Ala Pro Gly Ala Val His Leu 180 185 190	735
CCC CAG CCA GTG TCC ACA CGA TCC CAA CAC ACG CAG CCA ACT CCA GAA Pro Gln Pro Val Ser Thr Arg Ser Gln His Thr Gln Pro Thr Pro Glu 195 200 205 210	783
CCC AGC ACT GCT CCA AGC ACC TCC TTC CTG CTC CCA ATG GGC CCC AGC Pro Ser Thr Ala Pro Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser 215 220 225	831
CCC CCA GCT GAA GGG AGC ACT GGC GAC TTC GCT CTT CCA GTT GGA CTG Pro Pro Ala Glu Gly Ser Thr Gly Asp Phe Ala Leu Pro Val Gly Leu 230 235 240	879

SUBSTITUTE SHEET

2133326

ATT GTG GGT GTG ACA GCC TTG GGT CTA CTA ATA ATA GGA GTG GTG AAC	927
Ile Val Gly Val Thr Ala Leu Gly Leu Leu Ile Ile Gly Val Val Asn	
245 250 255	
TGT GTC ATC ATG ACC CAG GTG AAA AAG AAG CCC TTG TGC CTG CAG AGA	975
Cys Val Ile Met Thr Gln Val Lys Lys Lys Pro Leu Cys Leu Gln Arg	
260 265 270	
GAA GCC AAG GTG CCT CAC TTG CCT GCC GAT AAG GCC CGG GGT ACA CAG	1023
Glu Ala Lys Val Pro His Leu Pro Ala Asp Lys Ala Arg Gly Thr Gln	
275 280 285 290	
GGC CCC GAG CAG CAG CAC CTG CTG ATC ACA GCG CCG AGC TCC AGC AGC	1071
Gly Pro Glu Gln Gln His Leu Leu Ile Thr Ala Pro Ser Ser Ser Ser	
295 300 305	
AGC TCC CTG GAG AGC TCG GCC AGT GCG TTG GAC AGA AGG GCG CCC ACT	1119
Ser Ser Leu Glu Ser Ser Ala Ser Ala Leu Asp Arg Arg Ala Pro Thr	
310 315 320	
CGG AAC CAG CCA CAG GCA CCA GGC GTG GAG GCC AGT GGG GCC GGG GAG	1167
Arg Asn Gln Pro Gln Ala Pro Gly Val Glu Ala Ser Gly Ala Gly Glu	
325 330 335	
GCC CGG GCC AGC ACC GGG AGC TCA GAT TCT TCC CCT GGT GGC CAT GGG	1215
Ala Arg Ala Ser Thr Gly Ser Ser Asp Ser Ser Pro Gly Gly His Gly	
340 345 350	
ACC CAG GTC AAT GTC ACC TGC ATC GTG AAC GTC TGT AGC AGC TCT GAC	1263
Thr Gln Val Asn Val Thr Cys Ile Val Asn Val Cys Ser Ser Ser Asp	
355 360 365 370	
CAC AGC TCA CAG TGC TCC TCC CAA GCC AGC TCC ACA ATG GGA GAC ACA	1311
His Ser Ser Gln Cys Ser Ser Gln Ala Ser Ser Thr Met Gly Asp Thr	
375 380 385	
GAT TCC AGC CCC TCG GAG TCC CCG AAG GAC GAG CAG GTC CCC TTC TCC	1359
Asp Ser Ser Pro Ser Glu Ser Pro Lys Asp Glu Gln Val Pro Phe Ser	
390 395 400	
AAG GAG GAA TGT GCC TTT CGG TCA CAG CTG GAG ACG CCA GAG ACC CTG	1407
Lys Glu Glu Cys Ala Phe Arg Ser Gln Leu Glu Thr Pro Glu Thr Leu	
405 410 415	
CTG GGG AGC ACC GAA GAG AAG CCC CTG CCC CTT GGA GTG CCT GAT GCT	1455
Leu Gly Ser Thr Glu Glu Lys Pro Leu Pro Leu Gly Val Pro Asp Ala	
420 425 430	
GGG ATG AAG CCC AGT TAACCAGGCC GGTGTGGGCT GTGTCGTAGC CAAGGTGGGC	1510
Gly Met Lys Pro Ser	
435 440	
TGAGCCCTGG CAGGATGACC CTGCGAAGGG GCCCTGGTCC TTCCAGGCC CCACCACTAG	1570
GACTCTGAGG CTCTTTCTGG GCCAAGTTCC TCTAGTGCCC TCCACAGCCG CAGCCTCCCT	1630
CTGACCTGCA G	1641

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 461 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ala	Pro	Val	Ala	Val	Trp	Ala	Ala	Leu	Ala	Val	Gly	Leu	Glu	Leu	-22	-20	-15	-10
Trp	Ala	Ala	Ala	His	Ala	Leu	Pro	Ala	Gln	Val	Ala	Phe	Thr	Pro	Tyr	-5	1	5	10
Ala	Pro	Glu	Pro	Gly	Ser	Thr	Cys	Arg	Leu	Arg	Glu	Tyr	Tyr	Asp	Gln	15	20	25	
Thr	Ala	Gln	Met	Cys	Cys	Ser	Lys	Cys	Ser	Pro	Gly	Gln	His	Ala	Lys	30	35	40	
Val	Phe	Cys	Thr	Lys	Thr	Ser	Asp	Thr	Val	Cys	Asp	Ser	Cys	Glu	Asp	45	50	55	
Ser	Thr	Tyr	Thr	Gln	Leu	Trp	Asn	Trp	Val	Pro	Glu	Cys	Leu	Ser	Cys	60	65	70	
Gly	Ser	Arg	Cys	Ser	Ser	Asp	Gln	Val	Glu	Thr	Gln	Ala	Cys	Thr	Arg	75	80	85	90
Glu	Gln	Asn	Arg	Ile	Cys	Thr	Cys	Arg	Pro	Gly	Trp	Tyr	Cys	Ala	Leu	95	100	105	
Ser	Lys	Gln	Glu	Gly	Cys	Arg	Leu	Cys	Ala	Pro	Leu	Arg	Lys	Cys	Arg	110	115	120	
Pro	Gly	Phe	Gly	Val	Ala	Arg	Pro	Gly	Thr	Glu	Thr	Ser	Asp	Val	Val	125	130	135	
Cys	Lys	Pro	Cys	Ala	Pro	Gly	Thr	Phe	Ser	Asn	Thr	Thr	Ser	Ser	Thr	140	145	150	
Asp	Ile	Cys	Arg	Pro	His	Gln	Ile	Cys	Asn	Val	Val	Ala	Ile	Pro	Gly	155	160	165	170
Asn	Ala	Ser	Met	Asp	Ala	Val	Cys	Thr	Ser	Thr	Ser	Pro	Thr	Arg	Ser	175	180	185	
Met	Ala	Pro	Gly	Ala	Val	His	Leu	Pro	Gln	Pro	Val	Ser	Thr	Arg	Ser	190	195	200	
Gln	His	Thr	Gln	Pro	Thr	Pro	Glu	Pro	Ser	Thr	Ala	Pro	Ser	Thr	Ser	205	210	215	
Phe	Leu	Leu	Pro	Met	Gly	Pro	Ser	Pro	Pro	Ala	Glu	Gly	Ser	Thr	Gly	220	225	230	

2133326

56

Asp Phe Ala Leu Pro Val Gly Leu Ile Val Gly Val Thr Ala Leu Gly  
 235 240 245 250  
 Leu Leu Ile Ile Gly Val Val Asn Cys Val Ile Met Thr Gln Val Lys  
 255 260 265  
 Lys Lys Pro Leu Cys Leu Gln Arg Glu Ala Lys Val Pro His Leu Pro  
 270 275 280  
 Ala Asp Lys Ala Arg Gly Thr Gln Gly Pro Glu Gln Gln His Leu Leu  
 285 290 295  
 Ile Thr Ala Pro Ser Ser Ser Ser Ser Ser Leu Glu Ser Ser Ala Ser  
 300 305 310  
 Ala Leu Asp Arg Arg Ala Pro Thr Arg Asn Gln Pro Gln Ala Pro Gly  
 315 320 325 330  
 Val Glu Ala Ser Gly Ala Gly Glu Ala Arg Ala Ser Thr Gly Ser Ser  
 335 340 345  
 Asp Ser Ser Pro Gly Gly His Gly Thr Gln Val Asn Val Thr Cys Ile  
 350 355 360  
 Val Asn Val Cys Ser Ser Ser Asp His Ser Ser Gln Cys Ser Ser Gln  
 365 370 375  
 Ala Ser Ser Thr Met Gly Asp Thr Asp Ser Ser Pro Ser Glu Ser Pro  
 380 385 390  
 Lys Asp Glu Gln Val Pro Phe Ser Lys Glu Glu Cys Ala Phe Arg Ser  
 395 400 405 410  
 Gln Leu Glu Thr Pro Glu Thr Leu Leu Gly Ser Thr Glu Glu Lys Pro  
 415 420 425  
 Leu Pro Leu Gly Val Pro Asp Ala Gly Met Lys Pro Ser  
 430 435

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1368 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: HOMO SAPIENS

(vii) IMMEDIATE SOURCE:

- (B) CLONE: lambda-gt10-7ctnfbp

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..1366

(ix) FEATURE:

(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 121..1363

(ix) FEATURE:

(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 1..120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GGC CTC TCC ACC GTG CCT GAC CTG CTG CTG CCG CTG GTG CTC CTG	48
Met Gly Leu Ser Thr Val Pro Asp Leu Leu Leu Pro Leu Val Leu Leu	
-40 -35 -30 -25	
GAG CTG TTG GTG GGA ATA TAC CCC TCA GGG GTT ATT GGA CTG GTC CCT	96
Glu Leu Leu Val Gly Ile Tyr Pro Ser Gly Val Ile Gly Leu Val Pro	
-20 -15 -10	
CAC CTA GGG GAC AGG GAG AAG AGA GAT AGT GTG TGT CCC CAA GGA AAA	144
His Leu Gly Asp Arg Glu Lys Arg Asp Ser Val Cys Pro Gln Gly Lys	
-5 1 5	
TAT ATC CAC CCT CAA AAT AAT TCG ATT TGC TGT ACC AAG TGC CAC AAA	192
Tyr Ile His Pro Gln Asn Asn Ser Ile Cys Cys Thr Lys Cys His Lys	
10 15 20	
GGA ACC TAC TTG TAC AAT GAC TGT CCA GGC CCG GGG CAG GAT ACG GAC	240
Gly Thr Tyr Leu Tyr Asn Asp Cys Pro Gly Pro Gly Gln Asp Thr Asp	
25 30 35 40	
TGC AGG GAG TGT GAG AGC GGC TCC TTC ACC GCT TCA GAA AAC CAC CTC	288
Cys Arg Glu Cys Glu Ser Gly Ser Phe Thr Ala Ser Glu Asn His Leu	
45 50 55	
AGA CAC TGC CTC AGC TGC TCC AAA TGC CGA AAG GAA ATG GGT CAG GTG	336
Arg His Cys Leu Ser Cys Ser Lys Cys Arg Lys Glu Met Gly Gln Val	
60 65 70	
GAG ATC TCT TCT TGC ACA GTG GAC CGG GAC ACC GTG TGT GGC TGC AGG	384
Glu Ile Ser Ser Cys Thr Val Asp Arg Asp Thr Val Cys Gly Cys Arg	
75 80 85	
AAG AAC CAG TAC CGG CAT TAT TGG AGT GAA AAC CTT TTC CAG TGC TTC	432
Lys Asn Gln Tyr Arg His Tyr Trp Ser Glu Asn Leu Phe Gln Cys Phe	
90 95 100	
AAT TGC AGC CTC TGC CTC AAT GGG ACC GTG CAC CTC TCC TGC CAG GAG	480
Asn Cys Ser Leu Cys Leu Asn Gly Thr Val His Leu Ser Cys Gln Glu	
105 110 115 120	
AAA CAG AAC ACC GTG TGC ACC TGC CAT GCA GGT TTC TTT CTA AGA GAA	528
Lys Gln Asn Thr Val Cys Thr Cys His Ala Gly Phe Phe Leu Arg Glu	
125 130 135	

WO 93/19777  
2133326

PCT/US93/029

58

AAC Asn	GAG Glu	TGT Cys	GTC Val	TCC Ser	TGT Cys	AGT Ser	AAC Asn	TGT Cys	AAG Lys	AAA Lys	AGC Ser	CTG Leu	GAG Glu	TGC Cys	ACG Thr	576
			140					145					150			
AAG Lys	TTG Leu	TGC Cys	CTA Leu	CCC Pro	CAG Gln	ATT Ile	GAG Glu	AAT Asn	GTT Val	AAG Lys	GGC Gly	ACT Thr	GAG Glu	GAC Asp	TCA Ser	624
		155					160					165				
GGC Gly	ACC Thr	ACA Thr	GTG Val	CTG Leu	TTG Leu	CCC Pro	CTG Leu	GTC Val	ATT Ile	TTC Phe	TTT Phe	GGT Gly	CTT Leu	TGC Cys	CTT Leu	672
	170					175					180					
TTA Leu	TCC Ser	CTC Leu	CTC Leu	TTC Phe	ATT Ile	GGT Gly	TTA Leu	ATG Met	TAT Tyr	CGC Arg	TAC Tyr	CAA Gln	CGG Arg	TGG Trp	AAG Lys	720
					190					195					200	
TCC Ser	AAG Lys	CTC Leu	TAC Tyr	TCC Ser	ATT Ile	GTT Val	TGT Cys	GGG Gly	AAA Lys	TCG Ser	ACA Thr	CCT Pro	GAA Glu	AAA Lys	GAG Glu	768
				205					210						215	
GGG Gly	GAG Glu	CTT Leu	GAA Glu	GGA Gly	ACT Thr	ACT Thr	ACT Thr	AAG Lys	CCC Pro	CTG Leu	GCC Ala	CCA Pro	AAC Asn	CCA Pro	AGC Ser	816
			220					225							230	
TTC Phe	AGT Ser	CCC Pro	ACT Thr	CCA Pro	GGC Gly	TTC Phe	ACC Thr	CCC Pro	ACC Thr	CTG Leu	GGC Gly	TTC Phe	AGT Ser	CCC Pro	GTG Val	864
		235					240					245				
CCC Pro	AGT Ser	TCC Ser	ACC Thr	TTC Phe	ACC Thr	TCC Ser	AGC Ser	TCC Ser	ACC Thr	TAT Tyr	ACC Thr	CCC Pro	GGT Gly	GAC Asp	TGT Cys	912
	250					255					260					
CCC Pro	AAC Asn	TTT Phe	GCG Ala	GCT Ala	CCC Pro	CGC Arg	AGA Arg	GAG Glu	GTG Val	GCA Ala	CCA Pro	CCC Pro	TAT Tyr	CAG Gln	GGG Gly	960
	265				270					275					280	
GCT Ala	GAC Asp	CCC Pro	ATC Ile	CTT Leu	GCG Ala	ACA Thr	GCC Ala	CTC Leu	GCC Ala	TCC Ser	GAC Asp	CCC Pro	ATC Ile	CCC Pro	AAC Asn	1008
				285				290							295	
CCC Pro	CTT Leu	CAG Gln	AAG Lys	TGG Trp	GAG Glu	GAC Asp	AGC Ser	GCC Ala	CAC His	AAG Lys	CCA Pro	CAG Gln	AGC Ser	CTA Leu	GAC Asp	1056
			300					305							310	
ACT Thr	GAT Asp	GAC Asp	CCC Pro	GCG Ala	ACG Thr	CTG Leu	TAC Tyr	GCC Ala	GTG Val	GTG Val	GAG Glu	AAC Asn	GTG Val	CCC Pro	CCG Pro	1104
		315					320						325			
TTG Leu	CGC Arg	TGG Trp	AAG Lys	GAA Glu	TTC Phe	GTG Val	CGG Arg	CGC Arg	CTA Leu	GGG Gly	CTG Leu	AGC Ser	GAC Asp	CAC His	GAG Glu	1152
	330					335						340				
ATC Ile	GAT Asp	CGG Arg	CTG Leu	GAG Glu	CTG Leu	CAG Gln	AAC Asn	GGG Gly	CGC Arg	TGC Cys	CTG Leu	CGC Arg	GAG Glu	GCG Ala	CAA Gln	1200
	345				350					355					360	
TAC Tyr	AGC Ser	ATG Met	CTG Leu	GCG Ala	ACC Thr	TGG Trp	AGG Arg	CGG Arg	CGC Arg	ACG Thr	CCG Pro	CGG Arg	CGC Arg	GAG Glu	GCC Ala	1248
				365					370						375	

**SUBSTITUTE SHEET**

ACG CTG GAG CTG CTG GGA CGC GTG CTC CGC GAC ATG GAC CTG CTG GGC	1296
Thr Leu Glu Leu Leu Gly Arg Val Leu Arg Asp Met Asp Leu Leu Gly	
380 385 390	
TGC CTG GAG GAC ATC GAG GAG GCG CTT T GGC CCC GCC GCC CTC CCG	1344
Cys Leu Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro Ala Ala Leu Pro	
395 400 405	
CCC GCG CCC AGT CTT CTC AGA TGA	1368
Pro Ala Pro Ser Leu Leu Arg	
410 415	

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 455 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Leu Ser Thr Val Pro Asp Leu Leu Leu Pro Leu Val Leu Leu	-40	-35	-30	-25
Glu Leu Leu Val Gly Ile Tyr Pro Ser Gly Val Ile Gly Leu Val Pro	-20	-15	-10	
His Leu Gly Asp Arg Glu Lys Arg Asp Ser Val Cys Pro Gln Gly Lys	-5	1	5	
Tyr Ile His Pro Gln Asn Asn Ser Ile Cys Cys Thr Lys Cys His Lys	10	15	20	
Gly Thr Tyr Leu Tyr Asn Asp Cys Pro Gly Pro Gly Gln Asp Thr Asp	25	30	35	40
Cys Arg Glu Cys Glu Ser Gly Ser Phe Thr Ala Ser Glu Asn His Leu	45	50	55	
Arg His Cys Leu Ser Cys Ser Lys Cys Arg Lys Glu Met Gly Gln Val	60	65	70	
Glu Ile Ser Ser Cys Thr Val Asp Arg Asp Thr Val Cys Gly Cys Arg	75	80	85	
Lys Asn Gln Tyr Arg His Tyr Trp Ser Glu Asn Leu Phe Gln Cys Phe	90	95	100	
Asn Cys Ser Leu Cys Leu Asn Gly Thr Val His Leu Ser Cys Gln Glu	105	110	115	120
Lys Gln Asn Thr Val Cys Thr Cys His Ala Gly Phe Phe Leu Arg Glu	125	130	135	
Asn Glu Cys Val Ser Cys Ser Asn Cys Lys Lys Ser Leu Glu Cys Thr	140	145	150	



2133326

60

Lys Leu Cys Leu Pro Gln Ile Glu Asn Val Lys Gly Thr Glu Asp Ser  
 155 160 165  
 Gly Thr Thr Val Leu Leu Pro Leu Val Ile Phe Phe Gly Leu Cys Leu  
 170 175 180  
 Leu Ser Leu Leu Phe Ile Gly Leu Met Tyr Arg Tyr Gln Arg Trp Lys  
 185 190 195 200  
 Ser Lys Leu Tyr Ser Ile Val Cys Gly Lys Ser Thr Pro Glu Lys Glu  
 205 210 215  
 Gly Glu Leu Glu Gly Thr Thr Thr Lys Pro Leu Ala Pro Asn Pro Ser  
 220 225 230  
 Phe Ser Pro Thr Pro Gly Phe Thr Pro Thr Leu Gly Phe Ser Pro Val  
 235 240 245  
 Pro Ser Ser Thr Phe Thr Ser Ser Ser Thr Tyr Thr Pro Gly Asp Cys  
 250 255 260  
 Pro Asn Phe Ala Ala Pro Arg Arg Glu Val Ala Pro Pro Tyr Gln Gly  
 265 270 275 280  
 Ala Asp Pro Ile Leu Ala Thr Ala Leu Ala Ser Asp Pro Ile Pro Asn  
 285 290 295  
 Pro Leu Gln Lys Trp Glu Asp Ser Ala His Lys Pro Gln Ser Leu Asp  
 300 305 310  
 Thr Asp Asp Pro Ala Thr Leu Tyr Ala Val Val Glu Asn Val Pro Pro  
 315 320 325  
 Leu Arg Trp Lys Glu Phe Val Arg Arg Leu Gly Leu Ser Asp His Glu  
 330 335 340  
 Ile Asp Arg Leu Glu Leu Gln Asn Gly Arg Cys Leu Arg Glu Ala Gln  
 345 350 355 360  
 Tyr Ser Met Leu Ala Thr Trp Arg Arg Arg Thr Pro Arg Arg Glu Ala  
 365 370 375  
 Thr Leu Glu Leu Leu Gly Arg Val Leu Arg Asp Met Asp Leu Leu Gly  
 380 385 390  
 Cys Leu Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro Ala Ala Leu Pro  
 395 400 405  
 Pro Ala Pro Ser Leu Leu Arg  
 410 415

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3011 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:  
(B) CLONE: HU1L-1R
- (ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 84..1793
- (ix) FEATURE:  
(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 144..1790
- (ix) FEATURE:  
(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 84..143
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

AGACGCACCC TCTGAAGATG GTGGACTCCC TCCTGAGAAG CTGGGACCCC TTGGTAAAAG      60
ACAAGGCCTT CTCCAAGAAG AAT ATG AAA GTG TTA CTC AGA CTT ATT TGT      110
                Met Lys Val Leu Leu Arg Leu Ile Cys
                -20                               -15

TTC ATA GCT CTA CTG ATT TCT TCT CTG GAG GCT GAT AAA TGC AAG GAA      158
Phe Ile Ala Leu Leu Ile Ser Ser Leu Glu Ala Asp Lys Cys Lys Glu
-10                               -5                               1                               5

CGT GAA GAA AAA ATA ATT TTA GTG TCA TCT GCA AAT GAA ATT GAT GTT      206
Arg Glu Glu Lys Ile Ile Leu Val Ser Ser Ala Asn Glu Ile Asp Val
                10                               15                               20

CGT CCC TGT CCT CTT AAC CCA AAT GAA CAC AAA GGC ACT ATA ACT TGG      254
Arg Pro Cys Pro Leu Asn Pro Asn Glu His Lys Gly Thr Ile Thr Trp
                25                               30                               35

TAT AAA GAT GAC AGC AAG ACA CCT GTA TCT ACA GAA CAA GCC TCC AGG      302
Tyr Lys Asp Asp Ser Lys Thr Pro Val Ser Thr Glu Gln Ala Ser Arg
                40                               45                               50

ATT CAT CAA CAC AAA GAG AAA CTT TGG TTT GTT CCT GCT AAG GTG GAG      350
Ile His Gln His Lys Glu Lys Leu Trp Phe Val Pro Ala Lys Val Glu
                55                               60                               65

GAT TCA GGA CAT TAC TAT TGC GTG GTA AGA AAT TCA TCT TAC TGC CTC      398
Asp Ser Gly His Tyr Tyr Cys Val Val Arg Asn Ser Ser Tyr Cys Leu
                70                               75                               80                               85

AGA ATT AAA ATA AGT GCA AAA TTT GTG GAG AAT GAG CCT AAC TTA TGT      446
Arg Ile Lys Ile Ser Ala Lys Phe Val Glu Asn Glu Pro Asn Leu Cys
                90                               95                               100

```

2133326

TAT AAT GCA CAA GCC ATA TTT AAG CAG AAA CTA CCC GTT GCA GGA GAC	494
Tyr Asn Ala Gln Ala Ile Phe Lys Gln Lys Leu Pro Val Ala Gly Asp	
105 110 115	
GGA GGA CTT GTG TGC CCT TAT ATG GAG TTT TTT AAA AAT GAA AAT AAT	542
Gly Gly Leu Val Cys Pro Tyr Met Glu Phe Phe Lys Asn Glu Asn Asn	
120 125 130	
GAG TTA CCT AAA TTA CAG TGG TAT AAG GAT TGC AAA CCT CTA CTT CTT	590
Glu Leu Pro Lys Leu Gln Trp Tyr Lys Asp Cys Lys Pro Leu Leu Leu	
135 140 145	
GAC AAT ATA CAC TTT AGT GGA GTC AAA GAT AGG CTC ATC GTG ATG AAT	638
Asp Asn Ile His Phe Ser Gly Val Lys Asp Arg Leu Ile Val Met Asn	
150 155 160 165	
GTG GCT GAA AAG CAT AGA GGG AAC TAT ACT TGT CAT GCA TCC TAC ACA	686
Val Ala Glu Lys His Arg Gly Asn Tyr Thr Cys His Ala Ser Tyr Thr	
170 175 180	
TAC TTG GGC AAG CAA TAT CCT ATT ACC CGG GTA ATA GAA TTT ATT ACT	734
Tyr Leu Gly Lys Gln Tyr Pro Ile Thr Arg Val Ile Glu Phe Ile Thr	
185 190 195	
CTA GAG GAA AAC AAA CCC ACA AGG CCT GTG ATT GTG AGC CCA GCT AAT	782
Leu Glu Glu Asn Lys Pro Thr Arg Pro Val Ile Val Ser Pro Ala Asn	
200 205 210	
GAG ACA ATG GAA GTA GAC TTG GGA TCC CAG ATA CAA TTG ATC TGT AAT	830
Glu Thr Met Glu Val Asp Leu Gly Ser Gln Ile Gln Leu Ile Cys Asn	
215 220 225	
GTC ACC GGC CAG TTG AGT GAC ATT GCT TAC TGG AAG TGG AAT GGG TCA	878
Val Thr Gly Gln Leu Ser Asp Ile Ala Tyr Trp Lys Trp Asn Gly Ser	
230 235 240 245	
GTA ATT GAT GAA GAT GAC CCA GTG CTA GGG GAA GAC TAT TAC AGT GTG	926
Val Ile Asp Glu Asp Asp Pro Val Leu Gly Glu Asp Tyr Tyr Ser Val	
250 255 260	
GAA AAT CCT GCA AAC AAA AGA AGG AGT ACC CTC ATC ACA GTG CTT AAT	974
Glu Asn Pro Ala Asn Lys Arg Arg Ser Thr Leu Ile Thr Val Leu Asn	
265 270 275	
ATA TCG GAA ATT GAA AGT AGA TTT TAT AAA CAT CCA TTT ACC TGT TTT	1022
Ile Ser Glu Ile Glu Ser Arg Phe Tyr Lys His Pro Phe Thr Cys Phe	
280 285 290	
GCC AAG AAT ACA CAT GGT ATA GAT GCA GCA TAT ATC CAG TTA ATA TAT	1070
Ala Lys Asn Thr His Gly Ile Asp Ala Ala Tyr Ile Gln Leu Ile Tyr	
295 300 305	
CCA GTC ACT AAT TTC CAG AAG CAC ATG ATT GGT ATA TGT GTC ACG TTG	1118
Pro Val Thr Asn Phe Gln Lys His Met Ile Gly Ile Cys Val Thr Leu	
310 315 320 325	
ACA GTC ATA ATT GTG TGT TCT GTT TTC ATC TAT AAA ATC TTC AAG ATT	1166
Thr Val Ile Ile Val Cys Ser Val Phe Ile Tyr Lys Ile Phe Lys Ile	
330 335 340	

GAC ATT GTG CTT TGG TAC AGG GAT TCC TGC TAT GAT TTT CTC CCA ATA	1214
Asp Ile Val Leu Trp Tyr Arg Asp Ser Cys Tyr Asp Phe Leu Pro Ile	
345 350 355	
AAA GCT TCA GAT GGA AAG ACC TAT GAC GCA TAT ATA CTG TAT CCA AAG	1262
Lys Ala Ser Asp Gly Lys Thr Tyr Asp Ala Tyr Ile Leu Tyr Pro Lys	
360 365 370	
ACT GTT GGG GAA GGG TCT ACC TCT GAC TGT GAT ATT TTT GTG TTT AAA	1310
Thr Val Gly Glu Gly Ser Thr Ser Asp Cys Asp Ile Phe Val Phe Lys	
375 380 385	
GTC TTG CCT GAG GTC TTG GAA AAA CAG TGT GGA TAT AAG CTG TTC ATT	1358
Val Leu Pro Glu Val Leu Glu Lys Gln Cys Gly Tyr Lys Leu Phe Ile	
390 395 400 405	
TAT GGA AGG GAT GAC TAC GTT GGG GAA GAC ATT GTT GAG GTC ATT AAT	1406
Tyr Gly Arg Asp Asp Tyr Val Gly Glu Asp Ile Val Glu Val Ile Asn	
410 415 420	
GAA AAC GTA AAG AAA AGC AGA AGA CTG ATT ATC ATT TTA GTC AGA GAA	1454
Glu Asn Val Lys Lys Ser Arg Arg Leu Ile Ile Ile Leu Val Arg Glu	
425 430 435	
ACA TCA GGC TTC AGC TGG CTG GGT GGT TCA TCT GAA GAG CAA ATA GCC	1502
Thr Ser Gly Phe Ser Trp Leu Gly Gly Ser Ser Glu Glu Gln Ile Ala	
440 445 450	
ATG TAT AAT GCT CTT GTT CAG GAT GGA ATT AAA GTT GTC CTG CTT GAG	1550
Met Tyr Asn Ala Leu Val Gln Asp Gly Ile Lys Val Val Leu Leu Glu	
455 460 465	
CTG GAG AAA ATC CAA GAC TAT GAG AAA ATG CCA GAA TCG ATT AAA TTC	1598
Leu Glu Lys Ile Gln Asp Tyr Glu Lys Met Pro Glu Ser Ile Lys Phe	
470 475 480 485	
ATT AAG CAG AAA CAT GGG GCT ATC CGC TGG TCA GGG GAC TTT ACA CAG	1646
Ile Lys Gln Lys His Gly Ala Ile Arg Trp Ser Gly Asp Phe Thr Gln	
490 495 500	
GGA CCA CAG TCT GCA AAG ACA AGG TTC TGG AAG AAT GTC AGG TAC CAC	1694
Gly Pro Gln Ser Ala Lys Thr Arg Phe Trp Lys Asn Val Arg Tyr His	
505 510 515	
ATG CCA GTC CAG CGA CGG TCA CCT TCA TCT AAA CAC CAG TTA CTG TCA	1742
Met Pro Val Gln Arg Arg Ser Pro Ser Ser Lys His Gln Leu Leu Ser	
520 525 530	
CCA GCC ACT AAG GAG AAA CTG CAA AGA GAG GCT CAC GTG CCT CTC GGG	1790
Pro Ala Thr Lys Glu Lys Leu Gln Arg Glu Ala His Val Pro Leu Gly	
535 540 545	
TAGCATGGAG AAGTTGCCAA GAGTTCCTTA GGTGCCTCCT GTCTTATGGC GTTGCAGGCC	1850
AGGTTATGCC TCATGCTGAC TTGCAGAGTT CATGGAATGT AACTATATCA TCCTTTATCC	1910
CTGAGGTCAC CTGGAATCAG ATTATTAAGG GAATAAGCCA TGACGTCAAT AGCAGCCCAG	1970
GGCACTTCAG AGTAGAGGGC TTGGGAAGAT CTTTTAAAAA GGCAGTAGGC CCGGTGTGGT	2030

SUBSTITUTE SHEET

2133326

GGCTCACGCC TATAATCCCA GCACTTTGGG AGGCTGAAGT GGGTGGATCA CCAGAGGTCA 2090  
 GGAGTTCGAG ACCAGCCCAG CCAACATGGC AAAACCCCAT CTCTACTAAA AATACAAAAA 2150  
 TGAGCTAGGC ATGGTGGCAC ACGCCTGTAA TCCCAGCTAC ACCTGAGGCT GAGGCAGGAG 2210  
 AATTGCTTGA ACCGGGGAGA CGGAGGTTGC AGTGAGCCGA GTTTGGGCCA CTGCACTCTA 2270  
 GCCTGGCAAC AGAGCAAGAC TCCGTCTCAA AAAAAGGGCA ATAAATGCCC TCTCTGAATG 2330  
 TTTGAACTGC CAAGAAAAGG CATGGAGACA GCGAACTAGA AGAAAGGGCA AGAAGGAAAT 2390  
 AGCCACCGTC TACAGATGGC TTAGTTAAGT CATCCACAGC CCAAGGGCGG CGGCTATGCC 2450  
 TTGTCTGGGG ACCCTGTAGA GTCACTGACC CTGGAGCGGC TCTCCTGAGA GGTGCTGCAG 2510  
 GCAAAGTGAG ACTGACACCT CACTGAGGAA GGGAGACATA TTCTTGGAGA ACTTTCCATC 2570  
 TGCTTGTATT TTCCATACAC ATCCCCAGCC AGAAGTTAGT GTCCGAAGAA GAGCTTGAAA 2630  
 ACTCACTTCA ATGAACAAAG GGATTCTCCA GGATTCCAAA GTTTTGAAGT CATCTTAGCT 2690  
 TTCCACAGGA GGGAGAGAAC TAAAAAAGC AACAGTAGCA GGAATTGAT CCACTTCTTA 2750  
 ATGCTTTCCT CCCTGGCATG ACCATCCTGT CCTTTGTTAT TATCCTGCAT TTTACGTCTT 2810  
 TGGAGGAACA GCTCCCTAGT GGCTTCCTCC GTCTGCAATG TCCCTTGACAC AGCCCACACA 2870  
 TGAACCATCC TTCCCATGAT GCCGCTCTTC TGTCATCCCG CTCCTGCTGA AACACCTCCC 2930  
 AGGGGCTCCA CCTGTTCAGG AGCTGAAGCC CATGCTTTCC CACCAGCATG TCACTCCCAG 2990  
 ACCACCTCCC TGCCCTGTCC T 3011

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 569 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys Val Leu Leu Arg Leu Ile Cys Phe Ile Ala Leu Leu Ile Ser  
 -20 -15 -10 -5  
 Ser Leu Glu Ala Asp Lys Cys Lys Glu Arg Glu Glu Lys Ile Ile Leu  
 1 5 10  
 Val Ser Ser Ala Asn Glu Ile Asp Val Arg Pro Cys Pro Leu Asn Pro  
 15 20 25  
 Asn Glu His Lys Gly Thr Ile Thr Trp Tyr Lys Asp Asp Ser Lys Thr  
 30 35 40  
 Pro Val Ser Thr Glu Gln Ala Ser Arg Ile His Gln His Lys Glu Lys  
 45 50 55 60

Leu Trp Phe Val Pro Ala Lys Val Glu Asp Ser Gly His Tyr Tyr Cys  
65 70 75

Val Val Arg Asn Ser Ser Tyr Cys Leu Arg Ile Lys Ile Ser Ala Lys  
80 85 90

Phe Val Glu Asn Glu Pro Asn Leu Cys Tyr Asn Ala Gln Ala Ile Phe  
95 100 105

Lys Gln Lys Leu Pro Val Ala Gly Asp Gly Gly Leu Val Cys Pro Tyr  
110 115 120

Met Glu Phe Phe Lys Asn Glu Asn Asn Glu Leu Pro Lys Leu Gln Trp  
125 130 135 140

Tyr Lys Asp Cys Lys Pro Leu Leu Leu Asp Asn Ile His Phe Ser Gly  
145 150 155

Val Lys Asp Arg Leu Ile Val Met Asn Val Ala Glu Lys His Arg Gly  
160 165 170

Asn Tyr Thr Cys His Ala Ser Tyr Thr Tyr Leu Gly Lys Gln Tyr Pro  
175 180 185

Ile Thr Arg Val Ile Glu Phe Ile Thr Leu Glu Glu Asn Lys Pro Thr  
190 195 200

Arg Pro Val Ile Val Ser Pro Ala Asn Glu Thr Met Glu Val Asp Leu  
205 210 215 220

Gly Ser Gln Ile Gln Leu Ile Cys Asn Val Thr Gly Gln Leu Ser Asp  
225 230 235

Ile Ala Tyr Trp Lys Trp Asn Gly Ser Val Ile Asp Glu Asp Asp Pro  
240 245 250

Val Leu Gly Glu Asp Tyr Tyr Ser Val Glu Asn Pro Ala Asn Lys Arg  
255 260 265

Arg Ser Thr Leu Ile Thr Val Leu Asn Ile Ser Glu Ile Glu Ser Arg  
270 275 280

Phe Tyr Lys His Pro Phe Thr Cys Phe Ala Lys Asn Thr His Gly Ile  
285 290 295 300

Asp Ala Ala Tyr Ile Gln Leu Ile Tyr Pro Val Thr Asn Phe Gln Lys  
305 310 315

His Met Ile Gly Ile Cys Val Thr Leu Thr Val Ile Ile Val Cys Ser  
320 325 330

Val Phe Ile Tyr Lys Ile Phe Lys Ile Asp Ile Val Leu Trp Tyr Arg  
335 340 345

Asp Ser Cys Tyr Asp Phe Leu Pro Ile Lys Ala Ser Asp Gly Lys Thr  
350 355 360

Tyr Asp Ala Tyr Ile Leu Tyr Pro Lys Thr Val Gly Glu Gly Ser Thr  
365 370 375 380

2133326

66

Ser Asp Cys Asp Ile Phe Val Phe Lys Val Leu Pro Glu Val Leu Glu  
 385 390 395  
 Lys Gln Cys Gly Tyr Lys Leu Phe Ile Tyr Gly Arg Asp Asp Tyr Val  
 400 405 410  
 Gly Glu Asp Ile Val Glu Val Ile Asn Glu Asn Val Lys Lys Ser Arg  
 415 420 425  
 Arg Leu Ile Ile Ile Leu Val Arg Glu Thr Ser Gly Phe Ser Trp Leu  
 430 435 440  
 Gly Gly Ser Ser Glu Glu Gln Ile Ala Met Tyr Asn Ala Leu Val Gln  
 445 450 455 460  
 Asp Gly Ile Lys Val Val Leu Leu Glu Leu Glu Lys Ile Gln Asp Tyr  
 465 470 475  
 Glu Lys Met Pro Glu Ser Ile Lys Phe Ile Lys Gln Lys His Gly Ala  
 480 485 490  
 Ile Arg Trp Ser Gly Asp Phe Thr Gln Gly Pro Gln Ser Ala Lys Thr  
 495 500 505  
 Arg Phe Trp Lys Asn Val Arg Tyr His Met Pro Val Gln Arg Arg Ser  
 510 515 520  
 Pro Ser Ser Lys His Gln Leu Leu Ser Pro Ala Thr Lys Glu Lys Leu  
 525 530 535 540  
 Gln Arg Glu Ala His Val Pro Leu Gly  
 545

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1357 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (G) CELL TYPE: Human B cell lymphoblastoid
- (H) CELL LINE: CB23

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: pHuIL-1RII75

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 154..1350

(ix) FEATURE:  
 (A) NAME/KEY: mat\_peptide  
 (B) LOCATION: 193..1347

(ix) FEATURE:  
 (A) NAME/KEY: sig\_peptide  
 (B) LOCATION: 154..192

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGGAAAATA	CATTCTGCTA	CTCTTAAAAA	CTAGTGACGC	TCATACAAAT	CAACAGAAAG												60
AGCTTCTGAA	GGAAGACTTT	AAAGCTGCTT	CTGCCACGTG	CTGCTGGGTC	TCAGTCCTCC												120
ACTTCCCGTG	TCCTCTGGAA	GTTGTCAGGA	GCA	ATG	TTG	CGC	TTG	TAC	GTG	TTG							174
			Met	Leu	Arg	Leu	Tyr	Val	Leu								
			-13				-10										
GTA	ATG	GGA	GTT	TCT	GCC	TTC	ACC	CTT	CAG	CCT	GCG	GCA	CAC	ACA	GGG		222
Val	Met	Gly	Val	Ser	Ala	Phe	Thr	Leu	Gln	Pro	Ala	Ala	His	Thr	Gly		
	-5					1				5					10		
GCT	GCC	AGA	AGC	TGC	CGG	TTT	CGT	GGG	AGG	CAT	TAC	AAG	CGG	GAG	TTC		270
Ala	Ala	Arg	Ser	Cys	Arg	Phe	Arg	Gly	Arg	His	Tyr	Lys	Arg	Glu	Phe		
				15				20						25			
AGG	CTG	GAA	GGG	GAG	CCT	GTA	GCC	CTG	AGG	TGC	CCC	CAG	GTG	CCC	TAC		318
Arg	Leu	Glu	Gly	Glu	Pro	Val	Ala	Leu	Arg	Cys	Pro	Gln	Val	Pro	Tyr		
			30					35					40				
TGG	TTG	TGG	GCC	TCT	GTC	AGC	CCC	CGC	ATC	AAC	CTG	ACA	TGG	CAT	AAA		366
Trp	Leu	Trp	Ala	Ser	Val	Ser	Pro	Arg	Ile	Asn	Leu	Thr	Trp	His	Lys		
		45					50				55						
AAT	GAC	TCT	GCT	AGG	ACG	GTC	CCA	GGA	GAA	GAA	GAG	ACA	CGG	ATG	TGG		414
Asn	Asp	Ser	Ala	Arg	Thr	Val	Pro	Gly	Glu	Glu	Glu	Thr	Arg	Met	Trp		
	60					65					70						
GCC	CAG	GAC	GGT	GCT	CTG	TGG	CTT	CTG	CCA	GCC	TTG	CAG	GAG	GAC	TCT		462
Ala	Gln	Asp	Gly	Ala	Leu	Trp	Leu	Leu	Pro	Ala	Leu	Gln	Glu	Asp	Ser		
	75				80					85					90		
GGC	ACC	TAC	GTC	TGC	ACT	ACT	AGA	AAT	GCT	TCT	TAC	TGT	GAC	AAA	ATG		510
Gly	Thr	Tyr	Val	Cys	Thr	Thr	Arg	Asn	Ala	Ser	Tyr	Cys	Asp	Lys	Met		
				95				100						105			
TCC	ATT	GAG	CTC	AGA	GTT	TTT	GAG	AAT	ACA	GAT	GCT	TTC	CTG	CCG	TTC		558
Ser	Ile	Glu	Leu	Arg	Val	Phe	Glu	Asn	Thr	Asp	Ala	Phe	Leu	Pro	Phe		
			110					115					120				
ATC	TCA	TAC	CCG	CAA	ATT	TTA	ACC	TTG	TCA	ACC	TCT	GGG	GTA	TTA	GTA		606
Ile	Ser	Tyr	Pro	Gln	Ile	Leu	Thr	Leu	Ser	Thr	Ser	Gly	Val	Leu	Val		
		125					130					135					
TGC	CCT	GAC	CTG	AGT	GAA	TTC	ACC	CGT	GAC	AAA	ACT	GAC	GTG	AAG	ATT		654
Cys	Pro	Asp	Leu	Ser	Glu	Phe	Thr	Arg	Asp	Lys	Thr	Asp	Val	Lys	Ile		
	140					145					150						



2133326

CAA	TGG	TAC	AAG	GAT	TCT	CTT	CTT	TTG	GAT	AAA	GAC	AAT	GAG	AAA	TTT	702
Gln	Trp	Tyr	Lys	Asp	Ser	Leu	Leu	Leu	Asp	Lys	Asp	Asn	Glu	Lys	Phe	
155					160					165					170	
CTA	AGT	GTG	AGG	GGG	ACC	ACT	CAC	TTA	CTC	GTA	CAC	GAT	GTG	GCC	CTG	750
Leu	Ser	Val	Arg	Gly	Thr	Thr	His	Leu	Leu	Val	His	Asp	Val	Ala	Leu	
				175						180					185	
GAA	GAT	GCT	GGC	TAT	TAC	CGC	TGT	GTC	CTG	ACA	TTT	GCC	CAT	GAA	GGC	798
Glu	Asp	Ala	Gly	Tyr	Tyr	Arg	Cys	Val	Leu	Thr	Phe	Ala	His	Glu	Gly	
			190							195					200	
CAG	CAA	TAC	AAC	ATC	ACT	AGG	AGT	ATT	GAG	CTA	CGC	ATC	AAG	AAA	AAA	846
Gln	Gln	Tyr	Asn	Ile	Thr	Arg	Ser	Ile	Glu	Leu	Arg	Ile	Lys	Lys	Lys	
		205						210								
AAA	GAA	GAG	ACC	ATT	CCT	GTG	ATC	ATT	TCC	CCC	CTC	AAG	ACC	ATA	TCA	894
Lys	Glu	Glu	Thr	Ile	Pro	Val	Ile	Ile	Ser	Pro	Leu	Lys	Thr	Ile	Ser	
	220						225									
GCT	TCT	CTG	GGG	TCA	AGA	CTG	ACA	ATC	CCG	TGT	AAG	GTG	TTT	CTG	GGA	942
Ala	Ser	Leu	Gly	Ser	Arg	Leu	Thr	Ile	Pro	Cys	Lys	Val	Phe	Leu	Gly	
235						240									250	
ACC	GGC	ACA	CCC	TTA	ACC	ACC	ATG	CTG	TGG	TGG	ACG	GCC	AAT	GAC	ACC	990
Thr	Gly	Thr	Pro	Leu	Thr	Thr	Met	Leu	Trp	Trp	Thr	Ala	Asn	Asp	Thr	
					255										265	
CAC	ATA	GAG	AGC	GCC	TAC	CCG	GGA	GGC	CGC	GTG	ACC	GAG	GGG	CCA	CGC	1038
His	Ile	Glu	Ser	Ala	Tyr	Pro	Gly	Gly	Arg	Val	Thr	Glu	Gly	Pro	Arg	
			270						275						280	
CAG	GAA	TAT	TCA	GAA	AAT	AAT	GAG	AAC	TAC	ATT	GAA	GTG	CCA	TTG	ATT	1086
Gln	Glu	Tyr	Ser	Glu	Asn	Asn	Glu	Asn	Tyr	Ile	Glu	Val	Pro	Leu	Ile	
		285						290								
TTT	GAT	CCT	GTC	ACA	AGA	GAG	GAT	TTG	CAC	ATG	GAT	TTT	AAA	TGT	GTT	1134
Phe	Asp	Pro	Val	Thr	Arg	Glu	Asp	Leu	His	Met	Asp	Phe	Lys	Cys	Val	
	300							305								
GTC	CAT	AAT	ACC	CTG	AGT	TTT	CAG	ACA	CTA	CGC	ACC	ACA	GTC	AAG	GAA	1182
Val	His	Asn	Thr	Leu	Ser	Phe	Gln	Thr	Leu	Arg	Thr	Thr	Val	Lys	Glu	
315						320									330	
GCC	TCC	TCC	ACG	TTC	TCC	TGG	GGC	ATT	GTG	CTG	GCC	CCA	CTT	TCA	CTG	1230
Ala	Ser	Ser	Thr	Phe	Ser	Trp	Gly	Ile	Val	Leu	Ala	Pro	Leu	Ser	Leu	
					335										345	
GCC	TTC	TTG	GTT	TTG	GGG	GGA	ATA	TGG	ATG	CAC	AGA	CGG	TGC	AAA	CAC	1278
Ala	Phe	Leu	Val	Leu	Gly	Gly	Ile	Trp	Met	His	Arg	Arg	Cys	Lys	His	
				350						355					360	
AGA	ACT	GGA	AAA	GCA	GAT	GGT	CTG	ACT	GTG	CTA	TGG	CCT	CAT	CAT	CAA	1326
Arg	Thr	Gly	Lys	Ala	Asp	Gly	Leu	Thr	Val	Leu	Trp	Pro	His	His	Gln	
			365						370						375	
GAC	TTT	CAA	TCC	TAT	CCC	AAG	TGAAATAAAT									1357
Asp	Phe	Gln	Ser	Tyr	Pro	Lys										
	380					385										

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 398 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Leu	Arg	Leu	Tyr	Val	Leu	Val	Met	Gly	Val	Ser	Ala	Phe	Thr	Leu
-13			-10					-5						1	
Gln	Pro	Ala	Ala	His	Thr	Gly	Ala	Ala	Arg	Ser	Cys	Arg	Phe	Arg	Gly
	5					10					15				
Arg	His	Tyr	Lys	Arg	Glu	Phe	Arg	Leu	Glu	Gly	Glu	Pro	Val	Ala	Leu
20					25					30					35
Arg	Cys	Pro	Gln	Val	Pro	Tyr	Trp	Leu	Trp	Ala	Ser	Val	Ser	Pro	Arg
				40					45					50	
Ile	Asn	Leu	Thr	Trp	His	Lys	Asn	Asp	Ser	Ala	Arg	Thr	Val	Pro	Gly
			55					60					65		
Glu	Glu	Glu	Thr	Arg	Met	Trp	Ala	Gln	Asp	Gly	Ala	Leu	Trp	Leu	Leu
		70					75					80			
Pro	Ala	Leu	Gln	Glu	Asp	Ser	Gly	Thr	Tyr	Val	Cys	Thr	Thr	Arg	Asn
	85					90					95				
Ala	Ser	Tyr	Cys	Asp	Lys	Met	Ser	Ile	Glu	Leu	Arg	Val	Phe	Glu	Asn
100					105					110					115
Thr	Asp	Ala	Phe	Leu	Pro	Phe	Ile	Ser	Tyr	Pro	Gln	Ile	Leu	Thr	Leu
				120					125					130	
Ser	Thr	Ser	Gly	Val	Leu	Val	Cys	Pro	Asp	Leu	Ser	Glu	Phe	Thr	Arg
			135					140					145		
Asp	Lys	Thr	Asp	Val	Lys	Ile	Gln	Trp	Tyr	Lys	Asp	Ser	Leu	Leu	Leu
		150					155					160			
Asp	Lys	Asp	Asn	Glu	Lys	Phe	Leu	Ser	Val	Arg	Gly	Thr	Thr	His	Leu
		165				170					175				
Leu	Val	His	Asp	Val	Ala	Leu	Glu	Asp	Ala	Gly	Tyr	Tyr	Arg	Cys	Val
180					185					190					195
Leu	Thr	Phe	Ala	His	Glu	Gly	Gln	Gln	Tyr	Asn	Ile	Thr	Arg	Ser	Ile
				200					205					210	
Glu	Leu	Arg	Ile	Lys	Lys	Lys	Lys	Glu	Glu	Thr	Ile	Pro	Val	Ile	Ile
			215					220					225		
Ser	Pro	Leu	Lys	Thr	Ile	Ser	Ala	Ser	Leu	Gly	Ser	Arg	Leu	Thr	Ile
		230					235					240			

2133326

70

Pro Cys Lys Val Phe Leu Gly Thr Gly Thr Pro Leu Thr Thr Met Leu  
 245 250 255

Trp Trp Thr Ala Asn Asp Thr His Ile Glu Ser Ala Tyr Pro Gly Gly  
 260 265 270 275

Arg Val Thr Glu Gly Pro Arg Gln Glu Tyr Ser Glu Asn Asn Glu Asn  
 280 285 290

Tyr Ile Glu Val Pro Leu Ile Phe Asp Pro Val Thr Arg Glu Asp Leu  
 295 300 305

His Met Asp Phe Lys Cys Val Val His Asn Thr Leu Ser Phe Gln Thr  
 310 315 320

Leu Arg Thr Thr Val Lys Glu Ala Ser Ser Thr Phe Ser Trp Gly Ile  
 325 330 335

Val Leu Ala Pro Leu Ser Leu Ala Phe Leu Val Leu Gly Gly Ile Trp  
 340 345 350 355

Met His Arg Arg Cys Lys His Arg Thr Gly Lys Ala Asp Gly Leu Thr  
 360 365 370

Val Leu Trp Pro His His Gln Asp Phe Gln Ser Tyr Pro Lys  
 375 380 385

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGAAGGGAG CACTGGCGAC TAAGGATCCA

30

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 45 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GATCTGTAAC GTGGTGGCCA TCCCTGGGAA TGCAAGCATG GATGC

45

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGTCTGCACG TCCACGTCCC CCACCCGGTG AGC

33

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GATCTGTTGA GC

12

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTGAAACATC AGACGTGGTG TGCAAGCCCT GTTAAA

36

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGTCGACCT AGTGACGCTC ATACAAATC

29

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

2133326

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGCGGCCGC TCAGGAGGAG GCTTCCTTGA CTG

33

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..65

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CT GAA GGG AGC ACT GGC GAC GGT GGC GGT GGA TCC GGC GGT GGC GGC  
 Glu Gly Ser Thr Gly Asp Gly Gly Gly Gly Ser Gly Gly Gly Gly  
 1 5 10 15

47

GGC TCA TTG CCC GCC CAG G  
 Gly Ser Leu Pro Ala Gln  
 20

66

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Glu Gly Ser Thr Gly Asp Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly  
 1 5 10 15

Ser Leu Pro Ala Gln  
 20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ACCGAGGGAC CTGAGCG

17

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGAGCCGCCG CCACCGCCGG ATCCACCGCC ACCAGTGACT GGATATATTA ACT

53

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGTGGCGGTG GATCCGGCCG TGGCGGCGGC TCATTGCCCG CCCAGGTGGC A

51

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCTTGCTCAG CGCGCAGT

18

64166-33

74

CLAIMS:

1. A fusion protein comprising:

(1) a first tumor necrosis factor receptor (TNF-R) polypeptide covalently linked through a peptide linker to a second TNF-R polypeptide, or

(2) at least one TNF-R polypeptide covalently linked each through a peptide linker to at least one interleukin-1 receptor (IL-1R) polypeptide,

wherein each TNF-R polypeptide is a soluble TNF-R polypeptide and binds to tumor necrosis factor (TNF) and each IL-1R polypeptide is a soluble IL-1R polypeptide and binds to interleukin-1;

wherein in case of (1), the peptide linker comprises from 5 to 100 amino acids each selected from the group consisting of glycine, asparagine, serine, threonine and alanine; and

wherein in case of (2), each TNF-R polypeptide is human TNF-polypeptide and each IL-1R polypeptide is human IL-1R polypeptide.

2. A fusion protein in which at least one human tumor necrosis factor receptor (TNF-R) polypeptide and at least one human interleukin-1 receptor (IL-1R) polypeptide are covalently linked each through a peptide linker,

wherein each TNF-R is a soluble TNF-R and binds to tumor necrosis factor (TNF) and each IL-1R is a soluble IL-1R and binds to interleukin-1.

3. The fusion protein according to claim 2 represented by the formula:

64166-33

75

TNF-R - linker - TNF-R - linker - IL-1R, or

IL-1R - linker - TNF-R - linker - TNF-R,

wherein each linker is the peptide linker.

4. The fusion protein according to claim 3, wherein  
5 each of the peptide linkers comprises from 5 to 100 amino acids each selected from the group consisting of glycine, asparagine, serine, threonine, and alanine.
5. The fusion protein according to any one of  
10 claims 1 to 4, wherein each soluble TNF-R comprises an amino acid sequence selected from the group consisting of amino acids 1-x and -22-x of SEQ ID NO: 2, wherein x represents an integer from 163 to 235.
6. The fusion protein according to any one of  
15 claims 1 to 5, wherein each soluble IL-1R comprises an amino acid sequence selected from the group consisting of amino acids 1-x and -20-x of SEQ ID NO: 6, wherein x represents an integer from 312 to 316.
7. The fusion protein according to any one of  
20 claims 1 to 5, wherein each soluble IL-1R comprises an amino acid sequence selected from the group consisting of amino acids 1-x and -13-x of SEQ ID NO: 8, wherein x represents an integer from 330 to 333.
8. The fusion protein according to any one of  
claims 1 to 4,  
25 wherein each TNF-R is encoded by a TNF-R-encoding DNA selected from the group consisting of DNA comprising the nucleotide sequence presented in SEQ ID NO: 1 or SEQ ID NO: 3, and DNA capable of hybridizing under moderately stringent conditions to a complement of the



64166-33

76

nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, in which the TNF-R-encoding DNA encodes a biologically active TNF-R polypeptide; and

wherein each IL-1R is encoded by an IL-1R-encoding DNA selected from the group consisting of DNA comprising the nucleotide sequence presented in SEQ ID NO: 5 or SEQ ID NO: 7, and DNA capable of hybridizing under moderately stringent conditions to a complement of the nucleotide sequence of SEQ ID NO: 5 or SEQ ID NO: 7, in which the IL-1R-encoding DNA encodes a biologically active IL-1R polypeptide.

9. A fusion protein of the formula:

TNF-R - linker - TNF-R

wherein:

the linker is a peptide linker which comprises from 5 to 100 amino acids each selected from the group consisting of glycine, asparagine, serine, threonine, and alanine, and each TNF-R is a soluble tumor necrosis factor receptor (TNF-R) polypeptide which binds tumor necrosis factor (TNF).

10. The fusion protein according to claim 9, wherein each TNF-R is a human TNF-R.

11. The fusion protein according to claim 9, wherein each soluble TNF-R comprises an amino acid sequence selected from the group consisting of amino acids 1-x and -22-x of SEQ ID NO: 2, wherein x represents an integer from 163 to 235.

12. The fusion protein according to claim 9, wherein each TNF-R is encoded by a TNF-R-encoding DNA selected from

64166-33

77

the group consisting of DNA comprising the nucleotide sequence presented in SEQ ID NO: 1 or SEQ ID NO: 3, and DNA capable of hybridizing under moderately stringent conditions to a complement of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, in which the TNF-R-encoding DNA encodes a biologically active TNF-R polypeptide.

13. An isolated DNA sequence that encodes the fusion protein according to any one of claims 1 to 12.

14. An expression vector comprising the DNA sequence according to claim 13.

15. An isolated host cell transformed with the expression vector of claim 14.

16. A process for producing the fusion protein as defined in any one of claims 1 to 12, which comprises:

15 culturing a host cell transformed with the expression vector according to claim 14, under conditions that promote expression of the fusion protein, and recovering the fusion protein.

17. A pharmaceutical composition comprising:

20 the fusion protein according to any one of claims 1 to 12, and

a physiologically acceptable diluent, excipient, or carrier.

18. The pharmaceutical composition according to claim 17, which is for treating a condition selected from the group consisting of cachexia, rheumatoid arthritis,

64166-33

78

diabetes, multiple sclerosis, pulmonary fibrosis and silcosis, cerebral malaria, an allograft and xenograft rejection in graft versus host disease.

SMART & BIGGAR

OTTAWA, CANADA

PATENT AGENTS

Human

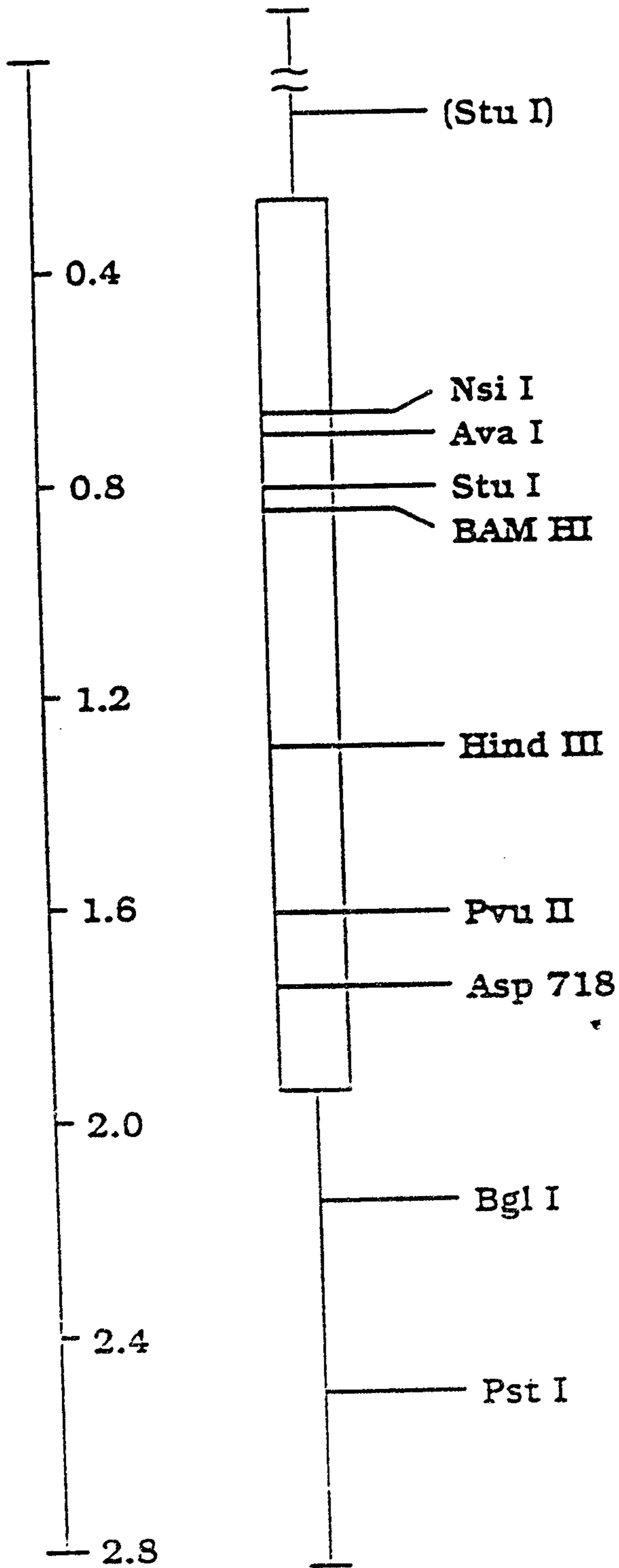


Figure 1



64166-33

3/6

## FIGURE 2 Cont'd.

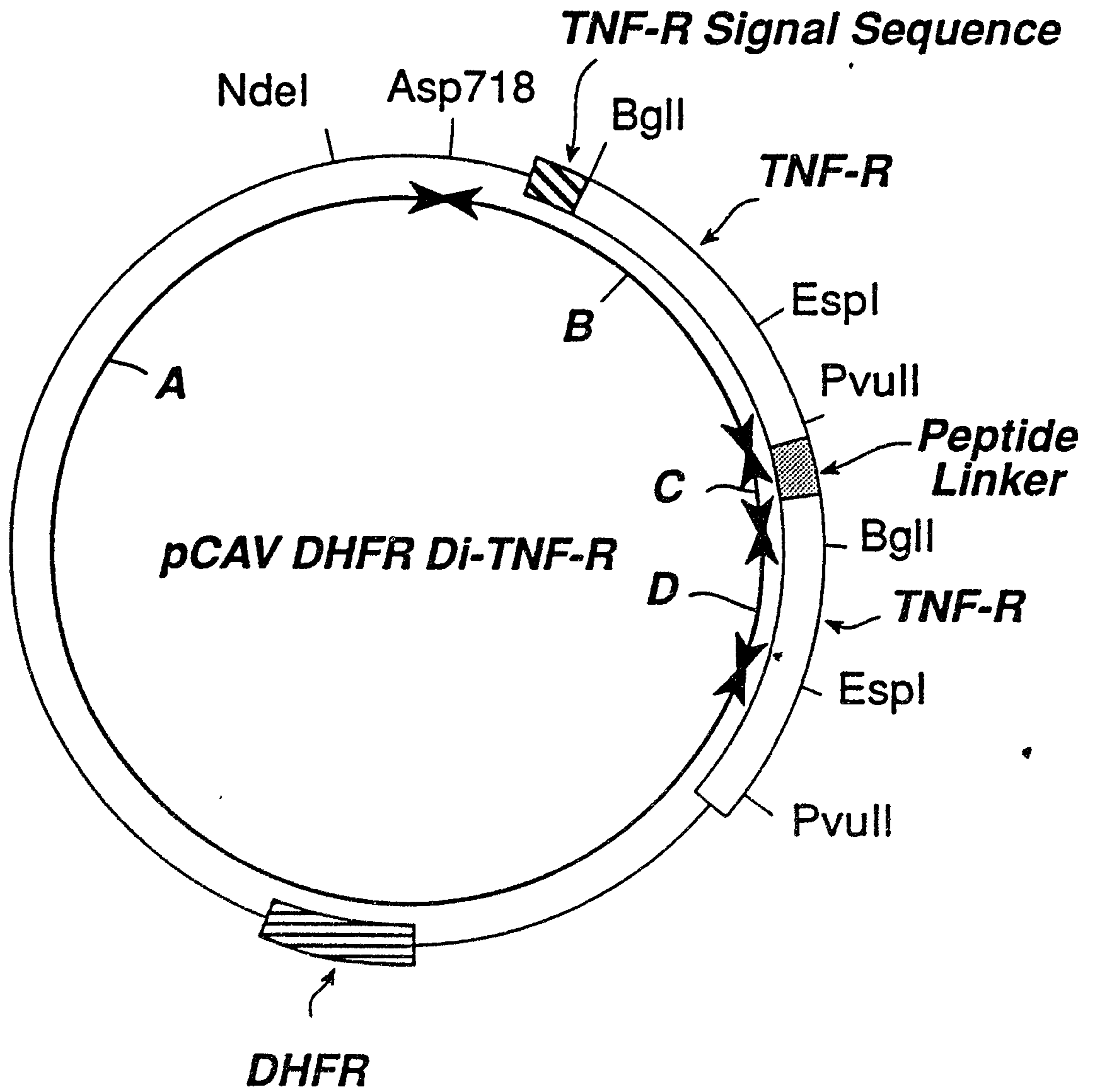
GGC	GAC	TTC	GCT	CTT	CCA	GTT	GGA	CTG	ATT	GTG	GGT	GTG	ACA	GCC	897
Gly	Asp	Phe	Ala	Leu	Pro	Val	Gly	Leu	Ile	Val	Gly	Val	Thr	Ala	248
TTG	GGT	CTA	CTA	ATA	ATA	GGA	GTG	GTG	AAC	TGT	GTC	ATC	ATG	ACC	942
Leu	Gly	Leu	Leu	Ile	Ile	Gly	Val	Val	Asn	Cys	Val	Ile	Met	Thr	263
CAG	GTG	AAA	AAG	AAG	CCC	TTG	TGC	CTG	CAG	AGA	GAA	GCC	AAG	GTG	987
Gln	Val	Lys	Lys	Lys	Pro	Leu	Cys	Leu	Gln	Arg	Glu	Ala	Lys	Val	278
CCT	CAC	TTG	CCT	GCC	GAT	AAG	GCC	CGG	GGT	ACA	CAG	GGC	CCC	GAG	1032
Pro	His	Leu	Pro	Ala	Asp	Lys	Ala	Arg	Gly	Thr	Gln	Gly	Pro	Glu	293
CAG	CAG	CAC	CTG	CTG	ATC	ACA	GCG	CCG	AGC	TCC	AGC	AGC	AGC	TCC	1077
Gln	Gln	His	Leu	Leu	Ile	Thr	Ala	Pro	Ser	Ser	Ser	Ser	Ser	Ser	308
CTG	GAG	AGC	TCG	GCC	AGT	GCG	TTG	GAC	AGA	AGG	GCG	CCC	ACT	CGG	1122
Leu	Glu	Ser	Ser	Ala	Ser	Ala	Leu	Asp	Arg	Arg	Ala	Pro	Thr	Arg	323
AAC	CAG	CCA	CAG	GCA	CCA	GGC	GTG	GAG	GCC	AGT	GGG	GCC	GGG	GAG	1167
Asn	Gln	Pro	Gln	Ala	Pro	Gly	Val	Glu	Ala	Ser	Gly	Ala	Gly	Glu	338
GCC	CGG	GCC	AGC	ACC	GGG	AGC	TCA	GAT	TCT	TCC	CCT	GGT	GGC	CAT	1212
Ala	Arg	Ala	Ser	Thr	Gly	Ser	Ser	Asp	Ser	Ser	Pro	Gly	Gly	His	353
GGG	ACC	CAG	GTC	AAT	GTC	ACC	TGC	ATC	GTG	AAC	GTC	TGT	AGC	AGC	1257
Gly	Thr	Gln	Val	Asn	Val	Thr	Cys	Ile	Val	Asn	Val	Cys	Ser	Ser	368
TCT	GAC	CAC	AGC	TCA	CAG	TGC	TCC	TCC	CAA	GCC	AGC	TCC	ACA	ATG	1302
Ser	Asp	His	Ser	Ser	Gln	Cys	Ser	Ser	Gln	Ala	Ser	Ser	Thr	Met	383
GGA	GAC	ACA	GAT	TCC	AGC	CCC	TCG	GAG	TCC	CCG	AAG	GAC	GAG	CAG	1347
Gly	Asp	Thr	Asp	Ser	Ser	Pro	Ser	Glu	Ser	Pro	Lys	Asp	Glu	Gln	398
GTC	CCC	TTC	TCC	AAG	GAG	GAA	TGT	GCC	TTT	CGG	TCA	CAG	CTG	GAG	1392
Val	Pro	Phe	Ser	Lys	Glu	Glu	Cys	Ala	Phe	Arg	Ser	Gln	Leu	Glu	413
ACG	CCA	GAG	ACC	CTG	CTG	GGG	AGC	ACC	GAA	GAG	AAG	CCC	CTG	CCC	1437
Thr	Pro	Glu	Thr	Leu	Leu	Gly	Ser	Thr	Glu	Glu	Lys	Pro	Leu	Pro	428
CTT	GGA	GTG	CCT	GAT	GCT	GGG	ATG	AAG	CCC	AGT					1470
Leu	Gly	Val	Pro	Asp	Ala	Gly	Met	Lys	Pro	Ser					439
TAACCAGGCCGGTGTGGGCTGTGTCGTAGCCAAGGTGGGCTGAGCCCTGGCAGGATGAC															
CCTGCGAAGGGGCCCTGGTCCTTCCAGGCCCCACCACTAGGACTCTGAGGCTCTTTCT															
GGGCCAAGTTCCTCTAGTGCCCTCCACAGCCGCGAGCCTCCCTCTGACCTGCAG....															

TTG	GGT	CTA	CTA	ATA	ATA	GGA	GTG	GTG	AAC	TGT	GTC	ATC	ATG	ACC	942
<u>Leu</u>	<u>Gly</u>	<u>Leu</u>	<u>Leu</u>	<u>Ile</u>	<u>Ile</u>	<u>Gly</u>	<u>Val</u>	<u>Val</u>	<u>Asn</u>	<u>Cys</u>	<u>Val</u>	<u>Ile</u>	<u>Met</u>	<u>Thr</u>	253
CAG	GTG	AAA	AAG	AAG	CCC	TTG	TGC	CTG	CAG	AGA	GAA	GCC	AAG	GTG	957
<u>Gln</u>	<u>Val</u>	Lys	Lys	Lys	Pro	Leu	Cys	Leu	Gln	Arg	Glu	Ala	Lys	Val	278
CCT	CAC	TTG	CCT	GCC	GAT	AAG	GCC	CGG	GGT	ACA	CAG	GGC	CCC	GAG	1032
Pro	His	Leu	Pro	Ala	Asp	Lys	Ala	Arg	Gly	Thr	Gln	Gly	Pro	Glu	293
CAG	CAG	CAC	CTG	CTG	ATC	ACA	GCG	CCG	AGC	TCC	AGC	AGC	AGC	TCC	1077
Gln	Gln	His	Leu	Leu	Ile	Thr	Ala	Pro	Ser	Ser	Ser	Ser	Ser	Ser	308
CTG	GAG	AGC	TCG	GCC	AGT	GCG	TTG	GAC	AGA	AGG	GCG	CCC	ACT	CGG	1122
Leu	Glu	Ser	Ser	Ala	Ser	Ala	Leu	Asp	Arg	Arg	Ala	Pro	Thr	Arg	323
AAC	CAG	CCA	CAG	GCA	CCA	GGC	GTG	GAG	GCC	AGT	GGG	GCC	GGG	GAG	1167
Asn	Gln	Pro	Gln	Ala	Pro	Gly	Val	Glu	Ala	Ser	Gly	Ala	Gly	Glu	338
GCC	CGG	GCC	AGC	ACC	GGG	AGC	TCA	GAT	TCT	TCC	CCT	GGT	GGC	CAT	1212
Ala	Arg	Ala	Ser	Thr	Gly	Ser	Ser	Asp	Ser	Ser	Pro	Gly	Gly	His	353
GGG	ACC	CAG	GTC	AAT	GTC	ACC	TGC	ATC	GTG	AAC	GTC	TGT	AGC	AGC	1257
Gly	Thr	Gln	Val	Asn	Val	Thr	Cys	Ile	Val	Asn	Val	Cys	Ser	Ser	368
TCT	GAC	CAC	AGC	TCA	CAG	TGC	TCC	TCC	CAA	GCC	AGC	TCC	ACA	ATG	1302
Ser	Asp	His	Ser	Ser	Gln	Cys	Ser	Ser	Gln	Ala	Ser	Ser	Thr	Met	383
GGA	GAC	ACA	GAT	TCC	AGC	CCC	TCG	GAG	TCC	CCG	AAG	GAC	GAG	CAG	1347
Gly	Asp	Thr	Asp	Ser	Ser	Pro	Ser	Glu	Ser	Pro	Lys	Asp	Glu	Gln	398
GTC	CCC	TTC	TCC	AAG	GAG	GAA	TGT	GCC	TTT	CGG	TCA	CAG	CTG	GAG	1392
Val	Pro	Phe	Ser	Lys	Glu	Glu	Cys	Ala	Phe	Arg	Ser	Gln	Leu	Glu	413
ACG	CCA	GAG	ACC	CTG	CTG	GGG	AGC	ACC	GAA	GAG	AAG	CCC	CTG	CCC	1437
Thr	Pro	Glu	Thr	Leu	Leu	Gly	Ser	Thr	Glu	Glu	Lys	Pro	Leu	Pro	428
CTT	GGA	GTG	CCT	GAT	GCT	GGG	ATG	AAG	CCC	AGT					1470
Leu	Gly	Val	Pro	Asp	Ala	Gly	Met	Lys	Pro	Ser					439
TAACCAGGCCGGTGTGGGCTGTGTCGTAGCCAAGGTGGGCTGAGCCCTGGCAGGATGAC															
CCTGCGAAGGGGCCCTGGTCCTTCCAGGCCCCCACTAGGACTCTGAGGCTCTTTCT															
GGGCCAAGTTCCTCTAGTGCCCTCCACAGCCGCAGCCTCCCTCTGACCTGCAG															

FIG. 3

4/6

Figure 3





5/6

Figure 4

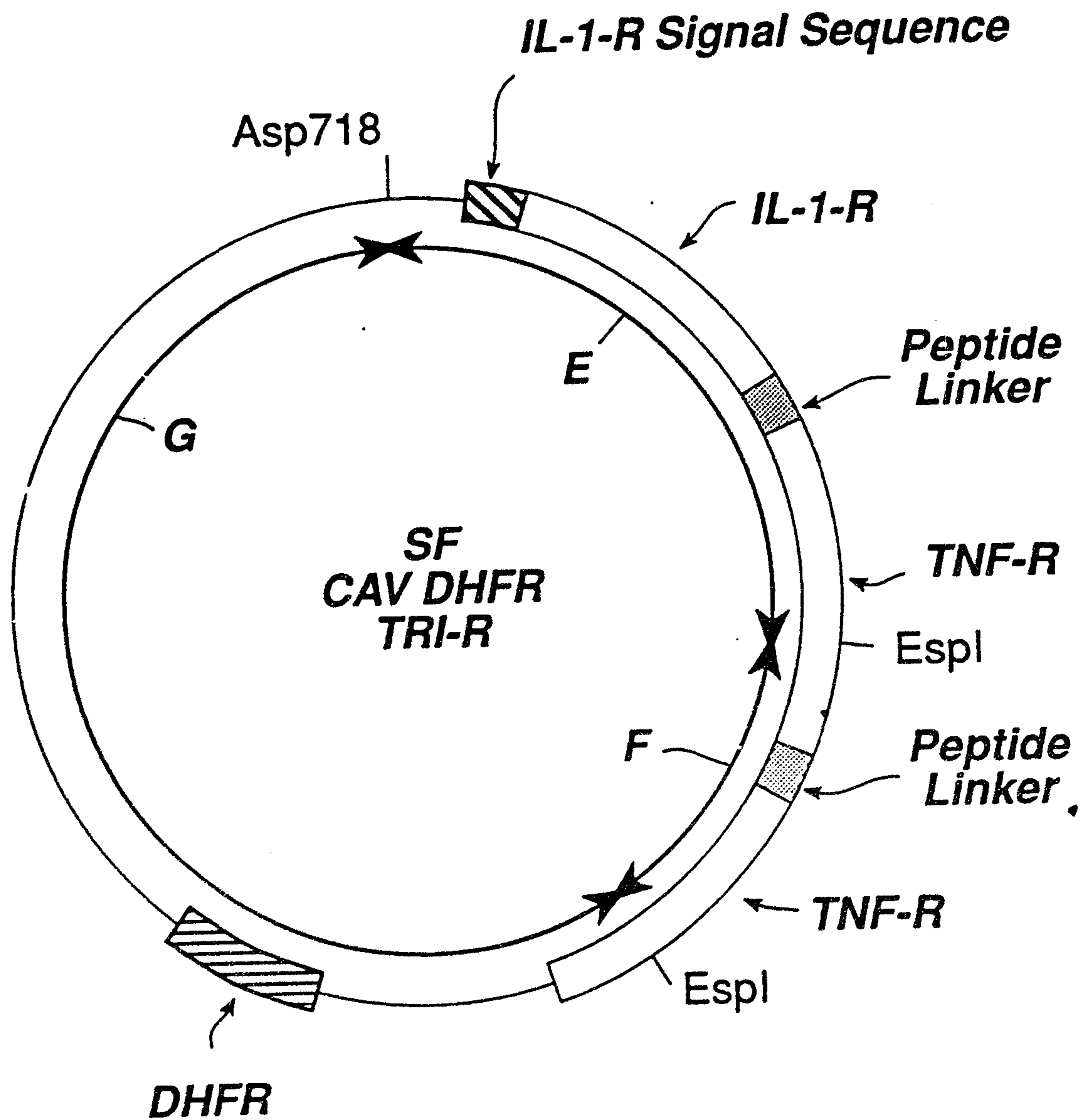


Figure 5

