US 20030165989A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2003/0165989 A1**

**Au-Young et al.** (43) **Pub. Date: Sep. 4, 2003**

(54) **GPCR DIAGNOSTIC FOR BRAIN CANCER**

(76) Inventors: **Janice Au-Young**, Brisbane, CA (US); **Karl J. Guegler**, Menlo Park, CA (US); **Muzong Cheng**, Oakland, CA (US)

Correspondence Address:
**INCYTE CORPORATION (formerly known as Incyte Genomics, Inc.)**
**3160 PORTER DRIVE**
**PALO ALTO, CA 94304 (US)**

(21) Appl. No.: **10/100,982**

(22) Filed: **Mar. 18, 2002**

**Related U.S. Application Data**

(60) Continuation-in-part of application No. 09/848,889, filed on May 3, 2001, now abandoned, which is a continuation-in-part of application No. 09/392,076, filed on Sep. 8, 1999, now abandoned, which is a division of application No. 08/812,871, filed on Mar. 6, 1997, now Pat. No. 5,955,303.

**Publication Classification**

(51) **Int. Cl.**[7] ..... **G01N 33/53**; C07H 21/04; C12P 21/04; C12P 21/02; C12N 5/06; C07K 14/715

(52) **U.S. Cl.** ..... **435/7.1**; 435/69.1; 435/70.21; 435/320.1; 435/325; 530/350; 536/23.5

(57) **ABSTRACT**

The invention provides a chemokine receptor-like protein, a cDNA encoding the protein and an antibody which specifically binds the protein. It also provides for the use of the cDNA, protein, and antibodies in the diagnosis, prognosis, treatment and evaluation of therapies for infection, inflammation and cancer, particularly meningioma of the brain. The invention further provides vectors and host cells for the production of the protein and transgenic model systems.

```
          9          18          27          36          45          54
5'CTA GTT CAA GAG GCC ATC TAC GAA CGT ATG ACT GCC GCT TTA AGA AGA CAG AGA

         63          72          81          90          99         108
GAA CTG AGT ATC CTC CCA AAG GTG ACA CTG GAA GCA ATG AAC ACC ACA GTG ATG
                                              M   N   T   T   V   M

        117         126         135         144         153         162
CAA GGC TTC AAC AGA TCT GAG CGG TGC CCC AGA GAC ACT CGG ATA GTA CAG CTG
Q   G   F   N   R   S   E   R   C   P   R   D   T   R   I   V   Q   L

        171         180         189         198         207         216
GTA TTC CCA GCC CTC TAC ACA GTG GTT TTC TTG ACC GGC ATC CTG CTG AAT ACT
V   F   P   A   L   Y   T   V   V   F   L   T   G   I   L   L   N   T

        225         234         243         252         261         270
TTG GCT CTG TGG GTG TTT GTT CAC ATC CCC AGC TCC TCC ACC TTC ATC ATC TAC
L   A   L   W   V   F   V   H   I   P   S   S   S   T   F   I   I   Y

        279         288         297         306         315         324
CTC AAA AAC ACT TTG GTG GCC GAC TTG ATA ATG ACA CTC ATG CTT CCT TTC AAA
L   K   N   T   L   V   A   D   L   I   M   T   L   M   L   P   F   K

        333         342         351         360         369         378
ATC CTC TCT GAC TCA CAC CTG GCA CCC TGG CAG CTC AGA GCT TTT GTG TGT CGT
I   L   S   D   S   H   L   A   P   W   Q   L   R   A   F   V   C   R

        387         396         405         414         423         432
TTT TCT TCG GTG ATA TTT TAT GAG ACC ATG TAT GTG GGC ATC GTG CTG TTA GGG
F   S   S   V   I   F   Y   E   T   M   Y   V   G   I   V   L   L   G

        441         450         459         468         477         486
CTC ATA GCC TTT GAC AGA TTC CTC AAG ATC ATC AGA CCT TTG AGA AAT ATT TTT
L   I   A   F   D   R   F   L   K   I   I   R   P   L   R   N   I   F

        495         504         513         522         531         540
CTA AAA AAA CCT GTT TTT GCA AAA ACG GTC TCA ATC TTC ATC TGG TTC TTT TTG
L   K   K   P   V   F   A   K   T   V   S   I   F   I   W   F   F   L

        549         558         567         576         585         594
TTC TTC ATC TCC CTG CCA ATT ATG ATC TTG AGC AAC AAG GAA GCA ACA CCA TCG
F   F   I   S   L   P   I   M   I   L   S   N   K   E   A   T   P   S

        603         612         621         630         639         648
TCT GTG AAA AAG TGT GCT TCC TTA AAG GGG CCT CTG GGG CTG AAA TGG CAT CAA
S   V   K   K   C   A   S   L   K   G   P   L   G   L   K   W   H   Q
```

FIGURE 1A

```
        657         666         675         684         693         702
ATG GTA AAT AAC ATA TGC CAG TTT ATT TTC TGG ACT GTT TTA ATC CTA ATG CTT
M   V   N   N   I   C   Q   F   I   F   W   T   V   L   I   L   M   L

        711         720         729         738         747         756
GTG TTT TAT GTG GTT ATT GCA AAA AAA GTA TAT GAT TCT TAT AGA AAG TCC AAA
V   F   Y   V   V   I   A   K   K   V   Y   D   S   Y   R   K   S   K

        765         774         783         792         801         810
TGT AAG GAC AGA AAA AAC AAC AAA AAG CTG GAA GGC AAA GTA TTT GTT GTC GTG
C   K   D   R   K   N   N   K   K   L   E   G   K   V   F   V   V   V

        819         828         837         846         855         864
CCT GTC TTC TTT GTG TGT TTT GCT CCA TTT CAT TTT GCC AGA GTT CCA TAT ACT
P   V   F   F   V   C   F   A   P   F   H   F   A   R   V   P   Y   T

        873         882         891         900         909         918
CAC AGT CAA ACC AAC AAT AAG ACT GAC TGT AGA CTG CAA AAT CAA CTG TTT ATT
H   S   Q   T   N   N   K   T   D   C   R   L   Q   N   Q   L   F   I

        927         936         945         954         963         972
GCT AAA GAA ACA ACT CTC TTT TTG GCA GCA ACT AAC ATT TGT ATG GAT CCC TTA
A   K   E   T   T   L   F   L   A   A   T   N   I   C   M   D   P   L

        981         990         999        1008        1017        1026
ATA TCC ATA TTC TTA TGT AAA AAA TTC ACA GAA AAG CTA CCA TGT ATG CAA GGG
I   S   I   F   L   C   K   K   F   T   E   K   L   P   C   M   Q   G

       1035        1044        1053        1062        1071        1080
AGA AAG ACC ACA GCA TCA AGC CAA GAA AAT CAT AGC AGT CAG ACA GAC AAC ATA
R   K   T   T   A   S   S   Q   E   N   H   S   S   Q   T   D   N   I

       1089        1098        1107        1116        1125        1134
ACC TTA GGC TGA CAA CTG TAC ATA GGG TTA ACT TCT ATT TAT TGA TGA GAC TTC
T   L   G

       1143        1152        1161        1170        1179        1188
CGT AGA TAA TGT GGA AAT CAA ATT TAA CCA AGA AAA AAA GAT TGG AAC AAA TGC

       1197        1206        1215        1224        1233        1242
TCT CTT ACA TTT TAT TAT CCT GGT GTA CAG AAA AGA TTA TAT AAA ATT TAA ATC

       1251        1260        1269        1278        1287        1296
CAC ATA GAT CTA TTC ATA AGC TGA ATG AAC CAT TAC TAA GAG AAT GCA ACA GGA
```

FIGURE 1B

```
      1305        1314        1323        1332        1341        1350
TAC AAA TGG CCA CTA GAG GTC ATT ATT TCT TTC TTT CTT TTT TTT TTT TTT TAA

      1359        1368        1377        1386        1395        1404
TTT CAA GAG CAT TTC ACT TTA ACA TTT TGG AAA AGA CTA AGG AGA AAC GTA TAT

      1413        1422        1431        1440        1449        1458
CCC TAC AAA CCT CCC CTC TAA ACA CCT TCT CAC ATT TTT TTC CAC AAT TCA CAT

      1467        1476        1485
AAC ACT ACT GCT TTT GTC CCC TTA AAT GT 3'
```

FIGURE 1C

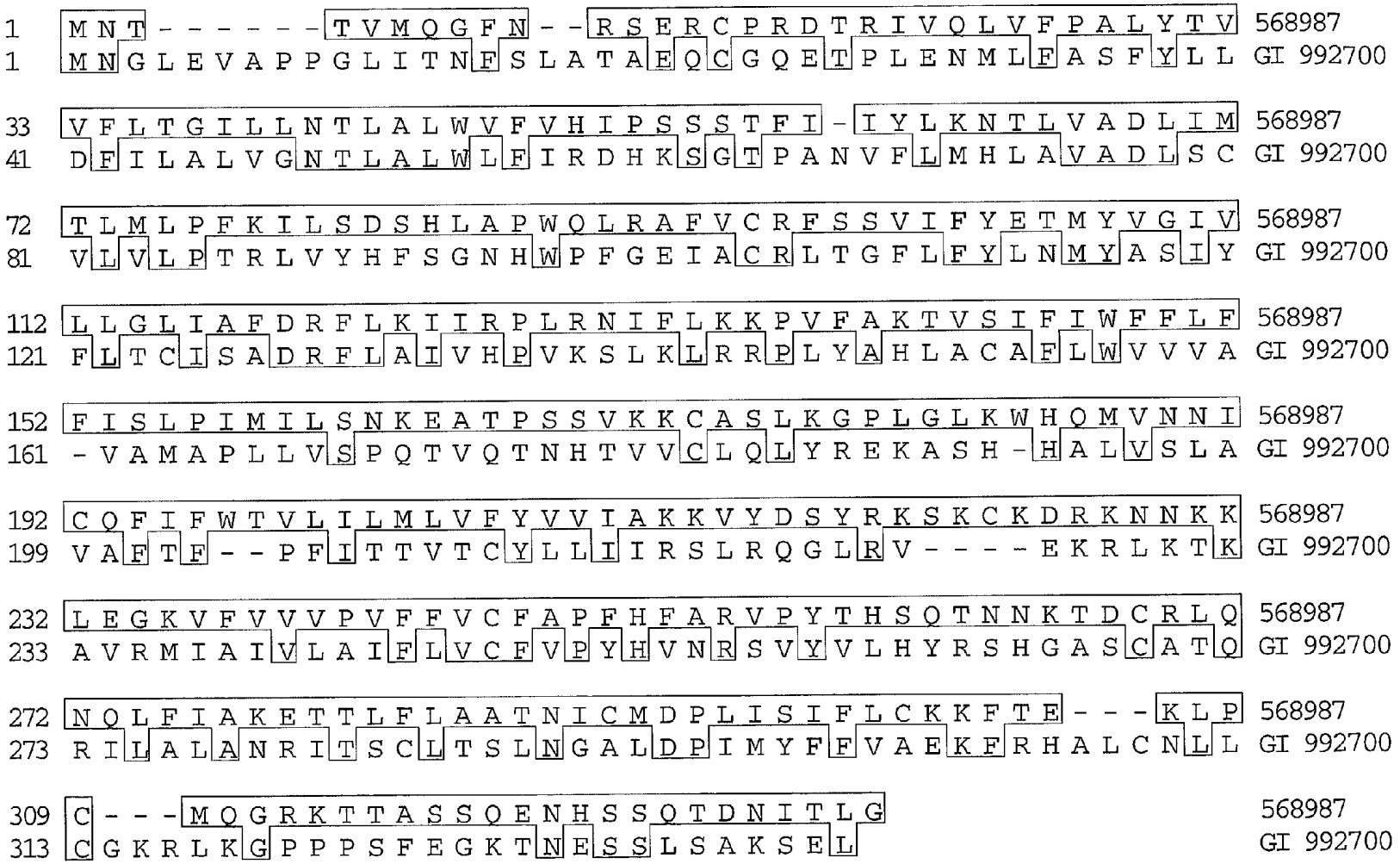
1 MNT------TVMQGFN--RSERCPRDTRIVQLVFPALYTV 568987
1 MNGLEVAPPGLITNFSLATAEQCGQETPLENMLFASFYLL GI 992700
33 VFLTGILLNTLALWVFVHIPSSSTFI-IYLKNTLVADLIM 568987
41 DFILALVGNTLALWLFIRDHKSGTPANVFLMHLAVADLSC GI 992700
72 TLMLPFKILSDSHLAPWQLRAFVCRFSSVIFYETMYVGIV 568987
81 VLVLPTRLVYHFSGNHWPFGEIACRLTGFLFYLNMYASIY GI 992700
112 LLGLIAFDRFLKIIRPLRNIFLKKPVFAKTVSIFIWFFLF 568987
121 FLTCISADRFLAIVHPVKSLKLRRPLYAHLACAFLWVVVA GI 992700
152 FISLPIMILSNKEATPSSVKKCASLKGPLGLKWHQMVNNI 568987
161 -VAMAPLLVSPQTVQTNHTVVCLQLYREKASH-HALVSLA GI 992700
192 CQFIFWTVLILMLVFYVVIAKKVYDSYRKSKCKDRKNNKK 568987
199 VAFTF--PFITTVTCYLLIIRSLRQGLRV----EKRLKTK GI 992700
232 LEGKVFVVVPVFFVCFAPFHFARVPYTHSQTNNKTDCRLQ 568987
233 AVRMIAIVLAIFLVCFVPYHVNRSVYVLHYRSHGASCATQ GI 992700
272 NQLFIAKETTLFLAATNICMDPLISIFLCKKFTE---KLP 568987
273 RILALANRITSCLTSLNGALDPIMYFFVAEKFRHALCNLL GI 992700
309 C---MQGRKTTASSQENHSSQTDNITLG 568987
313 CGKRLKGPPPSFEGKTNESSLSAKSEL GI 992700

FIGURE 2

| Category | cDNAs | Libraries | Abundance | % Abundance |
|---|---|---|---|---|
| Cardiovascular System | 266190 | 0/68 | 0 | 0.0000 |
| Connective Tissue | 144645 | 0/47 | 0 | 0.0000 |
| Digestive System | 501101 | 6/148 | 6 | 0.0012 |
| Embryonic Structures | 106713 | 0/21 | 0 | 0.0000 |
| Endocrine System | 225386 | 0/53 | 0 | 0.0009 |
| Exocrine Glands | 254635 | 2/64 | 2 | 0.0008 |
| Female Reproductive | 427284 | 2/106 | 2 | 0.0014 |
| Male Reproductive | 448207 | 1/114 | 1 | 0.0002 |
| Germ Cells | 38282 | 0/5 | 0 | 0.0078 |
| Hemic/Immune System | 680277 | 3/159 | 3 | 0.0012 |
| Liver | 109378 | 0/35 | 0 | 0.0000 |
| Musculoskeletal System | 159280 | 0/47 | 0 | 0.0000 |
| Nervous System | 955753 | 4/198 | 6 | 0.0013 |
| Pancreas | 110207 | 0/24 | 0 | 0.0000 |
| Respiratory System | 390086 | 2/93 | 2 | 0.0010 |
| Sense Organs | 19256 | 0/8 | 0 | 0.0000 |
| Skin | 72292 | 0/15 | 0 | 0.0000 |
| Stomatognathic System | 12923 | 0/10 | 0 | 0.0000 |
| Unclassified/Mixed | 120926 | 0/13 | 0 | 0.0017 |
| Urinary Tract | 279062 | 2/64 | 2 | 0.0007 |
| Totals | 5321883 | 22/1292 | 24 | 0.0000 |

Figure 3A

| Library | cDNAs | Description | Abundance | % Abundance |
|---|---|---|---|---|
| Nervous system | | | | |
| BRAITUT22 | 3700 | brain tumor, meningioma, 76F | 1 | 0.027 |
| BRAITUT02 | 13437 | brain tumor, mets hypernephroma, 58M | 3 | 0.022 |
| BRAUNOT01 | 8018 | brain, aw/CHF, 35M | 1 | 0.012 |
| NGANNOT01 | 3629 | neuroganglion tumor, ganglioneuroma, 9M | 1 | 0.007 |
| Exocrine Glands | | | | |
| BRSTDIT01 | 3433 | breast, aw/intraductal cancer, 48F | 1 | 0.029 |
| BRSTNOT28 | 3742 | breast, PF changes, 40F | 1 | 0.026 |
| Female Reproductive | | | | |
| OVARDIT04 | 3473 | ovary, dermoid cyst, 22F | 1 | 0.028 |
| OVARNOT10 | 3899 | ovary, aw/leiomyomata, 52F | 1 | 0.025 |
| Digestive System | | | | |
| COLANOT02 | 3979 | colon, ascending, CUC, 25F | 1 | 0.025 |
| COLXTDT01 | 4152 | colon, aw/leiomyomata, 37F | 1 | 0.024 |
| COLNTUP17 | 4361 | colon tumor, adenoCA, 3', CGAP | 1 | 0.022 |
| COLNNOT11 | 6717 | colon, mw/adenoCA, aw/node mets, 60M | 1 | 0.014 |
| SINTTUT01 | 2596 | sm intestine tumor, ileum, carcinoid, 42M | 1 | 0.038 |
| SINTBST01 | 5938 | sm intestine, ileum, Crohn's, 18F | 1 | 0.016 |
| Hemic/Immune | | | | |
| EOSIHET02 | 9312 | periph blood, hypereosinophilia, 48M | 1 | 0.010 |
| LEUKNOT03 | 3825 | white blood cells, 27F | 1 | 0.026 |
| SPLNNOT04 | 7794 | spleen, 2M | 1 | 0.012 |
| Respiratory System | | | | |
| LUNGNOT27 | 3961 | lung, 17F | 1 | 0.025 |
| LUNGTUT11 | 4204 | lung tumor, squamous cell CA, 57M | 1 | 0.023 |
| Male Reproductive | | | | |
| PROSNOT06 | 8829 | prostate, mw/adenoCA, 57M | 1 | 0.011 |
| Urinary tract | | | | |
| BLADNOT08 | 3826 | bladder, 11M | 1 | 0.026 |
| BLADNOT09 | 4125 | bladder, mw/TC CA, aw/node mets, 58M | 1 | 0.024 |

Figure 3B

# GPCR DIAGNOSTIC FOR BRAIN CANCER

[0001] This application is a continuation-in-part of U.S. Ser. No. 09/848,889, filed 3 May 2001, which is a continuation-in-part of U.S. Ser. No. 09/392,076, filed 8 Sep. 1999, which was a divisional of U.S. Pat. No. 5,955,303, issued 21 Sep. 1999, which matured from U.S. Ser. No. 08/812,871, filed 6 Mar. 1997.

## FIELD OF THE INVENTION

[0002] This invention relates to a human chemokine receptor-like protein, encoding cDNA, an antibody which specifically binds the protein, and to the use of these molecules in the diagnosis, prognosis, treatment and evaluation of therapies for infection, inflammation, and cancer.

## BACKGROUND OF THE INVENTION

[0003] Phylogenetic relationships among organisms have been demonstrated many times, and studies from a diversity of prokaryotic and eukaryotic organisms suggest a more or less gradual evolution of molecules, biochemical and physiological mechanisms, and metabolic pathways. Despite different evolutionary pressures, the proteins of nematode, fly, rat, and man have common chemical and structural features and generally perform the same cellular function. Comparisons of the nucleic acid and protein sequences from organisms where structure and/or function are known accelerate the investigation of human sequences and allow the development of model systems for testing diagnostic and therapeutic agents for human conditions, diseases, and disorders.

[0004] Immune response and cancer are characterized by continuous cell proliferation, inflammation, and cell death. Several molecular pathways have been linked to these activities, their development and progression. In addition, the analysis of the differential expression of key genes in any of these pathways may be diagnostically or prognostically important. For example, the analysis of cytokine levels is known to be useful as a prognostic indicator for distinguishing between various histologically-similar melanomas (Porter et al. (2001) Ann Surg Oncol 8:116-122).

[0005] Chemokines are a large family of low molecular weight, inducible, secreted, pro-inflammatory cytokines which are produced by various cell types. They have been divided into several subfamilies on the basis of the positions of their conserved cysteines. The CXC family includes interleukin-8 (IL-8), growth regulatory gene, neutrophil-activating peptide-2, and platelet factor 4 (PF-4). Although IL-8 and PF-4 are both polymorphonuclear chemoattractants, angiogenesis is stimulated by IL-8 and inhibited by PF-4. The CC family includes monocyte chemoattractant protein-1 (MCP-1), RANTES (regulated on activation, normal T cell-expressed and secreted), macrophage inflammatory proteins (MIP-1$\alpha$, MIP-1$\beta$), and eotaxin. MCP-1 is secreted by numerous cell types including endothelial, epithelial, and hematopoietic cells, and is a chemoattractant for monocytes and CD45RO+ lymphocytes (Proost (1996) Int J Clin Lab Res 26:211-223; Raport (1996) J Biol Chem 271:17161-17166).

[0006] Cells respond to cytokines and chemokines through G-protein-coupled receptors. These receptors are seven transmembrane molecules which transduce their signal through heterotrimeric GTP-binding proteins. Stimulation of the GTP-binding protein complex by activated receptor leads to the exchange of guanosine diphosphate for guanosine triphosphate and regulates the activity of effector molecules. The distinct classes of each of the subunits differ in activity and specificity and can elicit inhibitory or stimulatory responses. For example, when stimulation of the known cytokine receptors shows agonist-dependent inhibition of adenylyl cyclase and mobilization of intracellular calcium, the receptor is coupling to $G_{\alpha}i$ subunits (Myers et al. (1995) J Biol Chem 270:5786-5792).

[0007] The chemokine receptors display a range of sequence diversity. The known chemokine receptor protein sequence identities range from 22 to 40%. For example, the R12 receptor is most similar to the R20 orphan receptor (which has homology with the angiotensin receptor) and shows between 22 and 26% homology to characterized chemokine receptors including IL-8A and B, and MCP-1$\alpha$ and 1$\beta$ (Murphy (1994) Annu Rev Immunol 12:593-633; Raport et al. (1996) J Leuk Biol 59:18-23; and He et al. (1997) Nature 385:645-649). Chemokine receptors play a major role in the mobilization and activation of cells of the immune system, and certain receptors can respond to multiple ligands. Chemokine receptors have been implicated in the damage attributed to cytokines that occurs in the brains of Alzheimer's patients (Xia and Hyman (1999) J Neurovirol 5:32-41), and the CXCR4 receptor is widely known for its interactions between HIV-1, membrane fusion and viral entry. CXCR4 has been found to be expressed in fetal development and in adult brain, spinal cord, and bone marrow, and it has been implicated in tumorigenesis and was expressed in leukemias, Burkitt's lymphoma, and cancers of the brain, breast and uterus. CXCR4 was highly overexpressed in glioblastoma multiforme tumors (Sehgal et al. (1998) J Surg Oncol 69:239-48).

[0008] Cancer markers are of great importance in determining familial predisposition to cancers and in the early diagnosis and prognosis of various cancers. Two markers which gained widespread prominence as diagnostics in the past decade were PSA for prostate cancer and BRCAs 1 and 2 for breast cancer. Although these markers were originally named and employed in a tissue and disease specific manner, it is now known that BRCA expression is also upregulated in prostate cancer. Although GPCR expression levels are often very low under normal conditions Zweiger (2001, *Transducing the Genome*. McGraw Hill, San Francisco Calif.), it is common to see their elevation in specific tissues under disease conditions. Zweiger (supra) and Glavas et al. (2001, Proc Natl Acad Sci 6319-6324) discuss the correspondence between mRNA and protein expression. It is the tissue specific expression patterns of these various proteins that makes them useful as markers for clinical diagnosis and targets for immunotherapy.

[0009] The discovery of a new chemokine receptor-like protein, a cDNA which encodes it, and an antibody which specifically binds the protein satisfies a need in the art by providing compositions which are useful in the diagnosis, prognosis, treatment and evaluation of therapies for infection, inflammation, and cancer.

## SUMMARY OF THE INVENTION

[0010] The present invention is based on the discovery of a chemokine receptor-like protein and its encoding cDNA

which are overexpressed in brain cancer. These cDNA, protein and an antibody which specifically binds the protein are useful in the diagnosis, prognosis, treatment and evaluation of therapies for inflammation caused by bacterial infection and for cancer, particularly meningioma of the brain.

[0011] The invention provides an isolated cDNA comprising a nucleic acid sequence encoding a protein having the amino acid sequence of SEQ ID NO:1. The invention also provides an isolated cDNA selected from a nucleic acid sequence of SEQ ID NO:2, a fragment of SEQ ID NO:2 selected from SEQ ID NOs:3-10, and a variant of SEQ ID NO:2, SEQ ID NO:11 which has about 90% identity with SEQ ID NO:2, and the complements of SEQ ID NOs:2-11. The invention additionally provides compositions, a substrate, and a probe comprising the cDNA or the complement of the cDNA. The invention further provides a vector containing the cDNA, a host cell containing the vector and a method for using the cDNA to make the chemokine receptor-like protein. The invention still further provides a transgenic cell line or organism comprising the vector containing the cDNA encoding chemokine receptor-like protein. The invention additionally provides a fragment, a variant, or the complement of a cDNA selected from SEQ ID NOs:2-11. In one aspect, the invention provides a substrate containing at least one nucleotide sequence selected from SEQ ID NOs:2-11 or the complements thereof. In a second aspect, the invention provides a probe comprising a cDNA or the complement thereof which can be used in methods of detection, screening, and purification. In a further aspect, the probe is selected from a single-stranded RNA or DNA molecule, a peptide nucleic acid, a branched nucleic acid and the like.

[0012] The invention provides a method for using a cDNA to detect the differential expression of a nucleic acid in a sample comprising hybridizing a probe to the nucleic acids, thereby forming hybridization complexes and comparing hybridization complex formation with at least one standard, wherein the comparison confirms the differential expression of the cDNA in the sample. In one aspect, the method of detection further comprises amplifying the nucleic acids of the sample prior to hybridization. In another aspect, the method showing differential expression of the cDNA is used to diagnose infection, inflammation or cancer, particularly meningioma of the brain. In yet another aspect, the cDNA or a fragment or a variant or the complements thereof may comprise an element on an array.

[0013] The invention additionally provides a method for using a cDNA or a fragment or a variant or the complements thereof to screen a library or plurality of molecules or compounds to identify at least one ligand which specifically binds the cDNA, the method comprising combining the cDNA with the molecules or compounds under conditions allowing specific binding, and detecting specific binding to the cDNA , thereby identifying a ligand which specifically binds the cDNA. In one aspect, the molecules or compounds are selected from aptamers, DNA molecules, RNA molecules, peptide nucleic acids, artificial chromosome constructions, peptides, transcription factors, repressors, and regulatory molecules.

[0014] The invention provides a purified protein or a portion thereof selected from the group consisting of an amino acid sequence of SEQ ID NO:1, a variant of SEQ ID NO:1, an antigenic epitope of SEQ ID NO:1, and a biologically active portion of SEQ ID NO:1. The invention additionally provides an antigenic and biologically active portion of the protein of claim 2 consisting of residues V108 -I133. The invention also provides a composition comprising the purified protein. The invention further provides a method of using the chemokine receptor-like protein to treat a subject with infection, inflammation or cancer comprising administering to a patient in need of such treatment the composition containing the purified protein. The invention still further provides a method for using a protein to screen a library or a plurality of molecules or compounds to identify at least one ligand, the method comprising combining the protein with the molecules or compounds under conditions to allow specific binding and detecting specific binding, thereby identifying a ligand which specifically binds the protein. In one aspect, the molecules or compounds are selected from agonists, antagonists, DNA molecules, mimetics, peptides, peptide nucleic acids, proteins, and RNA molecules. In another aspect, the ligand is used to treat a subject with infection, inflammation, or cancer.

[0015] The invention provides a method for using a protein to identify an antibody that specifically binds the protein comprising contacting a plurality of antibodies with the protein under conditions to allow specific binding, and detecting specific binding between an antibody and the protein thereby identifying an antibody that specifically binds the protein. In one aspect, the antibody is selected from immunoglobulin molecules, polyclonal antibodies, monoclonal antibodies, chimeric antibodies, recombinant antibodies, humanized antibodies, single chain antibodies, Fab fragments, an $F(ab')_2$ fragments, Fv fragments; and antibody-peptide fusion proteins.

[0016] The invention provides an purified antibody that specifically binds the chemokine receptor-like protein. The invention also provides a composition comprising the purified antibody and a labeling moiety or a pharmaceutical agent. The invention additionally provides methods for using a protein to prepare and purify polyclonal and monoclonal antibodies that specifically bind the protein. The method for preparing a polyclonal antibody comprises immunizing a animal with protein under conditions to elicit an antibody response, isolating animal antibodies, attaching the protein to a substrate, contacting the substrate with isolated antibodies under conditions to allow specific binding to the protein, dissociating the antibodies from the protein, thereby obtaining polyclonal antibodies that specifically bind the protein. The invention further provides a polyclonal antibody that specifically binds the protein. The method for preparing a monoclonal antibody comprises immunizing a animal with a protein under conditions to elicit an antibody response, isolating antibody producing cells from the animal, fusing the antibody producing cells with immortalized cells in culture to form monoclonal antibody producing hybridoma cells, culturing the hybridoma cells, and isolating monoclonal antibodies from culture. The invention still further provides a monoclonal antibody that specifically binds the protein.

[0017] The invention provides a method for using an antibody that specifically binds the chemokine receptor-like protein to detect expression of a protein in a sample, the method comprising combining the antibody with a sample

under conditions for formation of antibody:protein complexes; and detecting complex formation, wherein complex formation indicates expression of the protein in the sample. In one aspect, the sample is brain tissue. In another aspect, the sample is peripheral blood. In yet another aspect, the amount of complex formation when compared to standards is diagnostic of brain cancer or Staphylococcus infection. In a still further aspect, the antibody is attached to a substrate. The invention also provides a method for using the antibody that specifically binds the protein in an assay to evaluate treatment of meningioma comprising contacting the antibody that specifically binds the chemokine receptor-like protein with a sample from a patient, detecting complex formation between the antibody and protein, comparing complex formation with standards, wherein the difference in complex formation indicates the efficacy of treatment.

[0018] The invention provides a method for immunopurification of a protein comprising attaching an antibody to a substrate, exposing the antibody to a sample containing protein under conditions to allow antibody:protein complexes to form, dissociating the protein from the complex, and collecting purified protein. The invention further provides a method of using an antibody to treat infection, inflammation and cancer comprising administering to a patient in need of such treatment a composition comprising the purified antibody.

[0019] The invention provides a method for inserting a heterologous marker gene into the genomic DNA of a mammal to disrupt the expression of the endogenous polynucleotide. The invention also provides a method for using a cDNA to produce a mammalian model system, the method comprising constructing a vector containing the cDNA of SEQ ID NO:11, transforming the vector into an embryonic stem cell, selecting a transformed embryonic stem cell, microinjecting the transformed embryonic stem cell into a mammalian blastocyst, thereby forming a chimeric blastocyst, transferring the chimeric blastocyst into a pseudopregnant dam, wherein the dam gives birth to a chimeric offspring containing the cDNA in its germ line, and breeding the chimeric mammal to produce a homozygous, mammalian model system.

## BRIEF DESCRIPTION OF THE FIGURES

[0020] **FIGS. 1A, 1B**, and 1C show the chemokine receptor-like protein, SEQ ID NO:1, encoded by the cDNA of SEQ ID NO:2. The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco Calif.).

[0021] **FIG. 2** demonstrates the conserved chemical and structural similarities among the chemokine receptor-like protein (568987; SEQ ID NO:1) and human chemokine receptor (g992700; SEQ ID NO:12). The alignment was produced using the MEGALIGN program of LASERGENE software (DNASTAR, Madison Wis.).

[0022] **FIGS. 3A and 3B** demonstrate northern analysis for the cDNA encoding the chemokine receptor-like protein. In **FIG. 3**A, the first column lists the category of cells or tissues; the second column, the number of cDNAs sequenced in that category; the third column, the number of libraries in which the cDNA is found versus the total number of libraries in that category; the fourth column, the abundance or number of cDNAs in that category; and the fifth column, the percent abundance (number of cDNAs divided by the total number of cDNAs in the category). In **FIG. 3**B, the first column lists the library name; the second column, the number of cDNAs sequenced for that library; the third column, the description of the tissue; the fourth column, abundance of the transcript; and the fifth column, percent abundance of the transcript.

## DESCRIPTION OF THE INVENTION

[0023] It is understood that this invention is not limited to the particular machines, materials and methods described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments and is not intended to limit the scope of the present invention which will be limited only by the appended claims. As used herein, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. For example, a reference to "a host cell" includes a plurality of such host cells known to those skilled in the art.

[0024] Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

[0025] Definitions

[0026] "Chemokine receptor-like protein" refers to a purified protein obtained from any mammalian species, including bovine, canine, murine, ovine, porcine, rodent, simian, and preferably the human species, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

[0027] "Antibody" refers to intact immunoglobulin molecule, a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a recombinant antibody, a humanized antibody, single chain antibodies, a Fab fragment, an $F(ab')_2$ fragment, an Fv fragment; and an antibody-peptide fusion protein.

[0028] "Antigenic determinant" refers to an antigenic or immunogenic epitope, structural feature, or region of an oligopeptide, peptide, or protein which is capable of inducing formation of an antibody which specifically binds the protein. Biological activity is not a prerequisite for immunogenicity.

[0029] "Array" refers to an ordered arrangement of at least two cDNAs, proteins, or antibodies on a substrate. At least one of the cDNAs, proteins, or antibodies represents a control or standard, and the other cDNA, protein, or antibody of diagnostic or therapeutic interest. The arrangement of at least two and up to about 40,000 cDNAs, proteins, or antibodies on the substrate assures that the size and signal intensity of each labeled complex, formed between each cDNA and at least one nucleic acid, each protein and at least one ligand or antibody, or each antibody and at least one protein to which the antibody specifically binds, is individually distinguishable.

[0030] The "complement" of a cDNA of the Sequence Listing refers to a nucleic acid molecule which is completely

complementary over its full length and which will hybridize to the cDNA or an mRNA under conditions of high stringency.

[0031] "cDNA" refers to an isolated polynucleotide, nucleic acid molecule, or any fragment thereof that contains from about 400 to about 12,000 nucleotides. It may have originated recombinantly or synthetically, may be double-stranded or single-stranded, represents coding and noncoding 3' or 5' sequence, and generally lacks introns.

[0032] The phrase "cDNA encoding a protein" refers to a nucleic acid whose sequence closely aligns with sequences that encode conserved regions, motifs or domains identified by employing analyses well known in the art. These analyses include BLAST (Basic Local Alignment Search Tool; Altschul (1993) J Mol Evol 36:290-300; Altschul et al. (1990) J Mol Biol 215:403-410) and BLAST2 (Altschul et al. (1997) Nucleic Acids Res 25:3389-3402) which provide identity within the conserved region. Brenner et al. (1998; Proc Natl Acad Sci 95:6073-6078) who analyzed BLAST for its ability to identify structural homologs by sequence identity found 30% identity is a reliable threshold for sequence alignments of at least 150 residues and 40% is a reasonable threshold for alignments of at least 70 residues (Brenner, page 6076, column 2).

[0033] A "composition" refers to the polynucleotide and a labeling moiety, a purified protein and a pharmaceutical carrier, an antibody and a labeling moiety, and the like.

[0034] "Derivative" refers to a cDNA or a protein that has been subjected to a chemical modification. Derivatization of a cDNA can involve substitution of a nontraditional base such as queosine or of an analog such as hypoxanthine. Derivatization of a protein involves the replacement of a hydrogen by an acetyl, acyl, alkyl, amino, formyl, or morpholino group. Derivative molecules retain the biological activities of the naturally occurring molecules but may confer advantages such as longer lifespan or enhanced activity.

[0035] "Differential expression" refers to an increased or upregulated or a decreased or downregulated expression as detected by absence, presence, or at least two-fold change in the amount of transcribed messenger RNA or translated protein in a sample.

[0036] "Disorder" refers to conditions, diseases or syndromes in which the cDNAs and chemokine receptor-like protein are differentially expressed. Such a disorder includes infection, particularly complications of bacterial and viral infection; inflammation, particularly chronic ulcerative colitis, Crohn's disease, or complications of cancer; and cancers, particularly adenocarcinomas of the colon and prostate, brain tumors (meningioma, hypernephroma), breast tumors (ductal or intraductal), neuroganglion tumors (ganglioneuroma), small intestine tumors (carcinoid), transitional cell carcinoma of the bladder, and leiomyomata of the uterus.

[0037] An "expression profile" is a representation of gene expression in a sample. A nucleic acid expression profile is produced using sequencing, hybridization, or amplification technologies and mRNAs or cDNAs from a sample. A protein expression profile, although time delayed, mirrors the nucleic acid expression profile and uses labeling moieties or antibodies to detect expression in a sample. The nucleic acids, proteins, or antibodies may be used in solution or attached to a substrate, and their detection is based on methods and labeling moieties well known in the art.

[0038] A "hybridization complex" is formed between a cDNA and a nucleic acid of a sample when the purines of one molecule hydrogen bond with the pyrimidines of the complementary molecule, e.g., 5'-A-G-T-C-3' base pairs with 3'-T-C-A-G-5'. Hybridization conditions, degree of complementarity and the use of nucleotide analogs affect the efficiency and stringency of hybridization reactions.

[0039] "Identity" as applied to sequences, refers to the quantification (usually percentage) of nucleotide or residue matches between at least two sequences aligned using a standardized algorithm such as Smith-Waterman alignment (Smith and Waterman (1981) J Mol Biol 147:195-197), CLUSTALW (Thompson et al. (1994) Nucleic Acids Res 22:4673-4680), or BLAST2 (Altschul et al. (1997) Nucleic Acids Res 25:3389-3402. BLAST2 may be used in a standardized and reproducible way to insert gaps in one of the sequences in order to optimize alignment and to achieve a more meaningful comparison between them. "Similarity" uses the same algorithms but takes conservative substitution of nucleotides and residues into account. In proteins, similarity exceeds identity in that substitution of a valine for a leucine or isoleucine, is counted in calculating the reported percentage. Substitutions which are considered to be conservative are well known in the art.

[0040] Isolated or "purified" refers to any molecule or compound that is separated from its natural environment and is from about 60% free to about 90% free from other components with which it is naturally associated.

[0041] "Labeling moiety" refers to any reporter molecule including radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, substrates, cofactors, inhibitors, or magnetic particles that can be attached to or incorporated into a polynucleotide, protein or antibody. Visible labels include but are not limited to anthocyanins, green fluorescent protein (GFP), β glucuronidase, luciferase, Cy3 and Cy5, and the like. Radioactive markers include radioactive forms of hydrogen, iodine, phosphorous, sulfur, and the like.

[0042] "Ligand" refers to any agent, molecule, or compound which will bind specifically to a polynucleotide or to an epitope of a protein. Such ligands stabilize or modulate the activity of polynucleotides or proteins and may be composed of inorganic and/or organic substances including minerals, cofactors, nucleic acids, proteins, carbohydrates, fats, and lipids.

[0043] An "oligopeptide" is an amino acid sequence from about five residues to about 15 residues that is used as part of a fusion protein to produce an antibody.

[0044] "Post-translational modification" of a protein can involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and the like. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cellular location, cell type, pH, enzymatic milieu, and the like.

[0045] "Protein" refers to a polypeptide or any portion thereof. A "portion" of a protein refers to that length of amino acid sequence which would retain at least one biological activity, a domain identified by PFAM or PRINTS

analysis or an antigenic determinant of the protein identified using Kyte-Doolittle algorithms of the PROTEAN program (DNASTAR, Madison Wis.).

[0046] "Sample" is used in its broadest sense as containing nucleic acids, proteins, antibodies, and the like. A sample may comprise a bodily fluid; the soluble fraction of a cell preparation, or an aliquot of media in which cells were grown; a chromosome, an organelle, or membrane isolated or extracted from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; a cell; a tissue; a tissue print; buccal cells, skin, hair or hair follicle; and the like.

[0047] "Specific binding" refers to a special and precise interaction between two molecules which is dependent upon their structure, particularly their molecular side groups. For example, the intercalation of a regulatory protein into the major groove of a DNA molecule or the binding between an epitope of a protein and an agonist, antagonist, or antibody.

[0048] "Substrate" refers to any rigid or semi-rigid support to which polynucleotides, proteins, or antibodies are bound and includes membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, capillaries or other tubing, plates, polymers, and microparticles with a variety of surface forms including wells, trenches, pins, channels and pores.

[0049] A "transcript image" (TI) is a profile of gene transcription activity in a particular tissue at a particular time. TI provides assessment of the relative abundance of expressed polynucleotides in the cDNA libraries of an EST database as described in U.S. Pat. No. 5,840,484, incorporated herein by reference.

[0050] "Variant" refers to molecules that are recognized variations of a cDNA or a protein encoded by the cDNA. Splice variants may be determined by BLAST score, wherein the score is at least 100, and most preferably at least 400. Allelic variants have a high percent identity to the cDNAs and may differ by about three bases per hundred bases. "Single nucleotide polymorphism" (SNP) refers to a change in a single base as a result of a substitution, insertion or deletion. The change may be conservative (purine for purine) or non-conservative (purine to pyrimidine) and may or may not result in a change in an encoded amino acid or its secondary, tertiary, or quaternary structure.

## THE INVENTION

[0051] The invention is based on the discovery of a cDNA which encodes a chemokine receptor-like protein and on the use of a cDNA, or fragments thereof; a protein, or portions thereof; or an antibody which specifically binds the protein directly or as compositions in the characterization, diagnosis, prognosis, treatment and evaluation of therapies for the inflammation associated with bacterial infection and for cancer, particularly meningioma of the brain.

[0052] Nucleic acids encoding the human chemokine receptor-like protein shown in **FIGS. 1A, 1B**, and 1C were first identified in Incyte Clone 568987 from the macrophage cDNA library, MMLR3DT01, through a computer-generated search for amino acid sequence alignments. The complete nucleotide sequence, SEQ ID NO:2, was derived from extension of Incyte clone 568987. It's nucleotide sequence has been confirmed by assembly of sequence fragments found in Incyte clones 1881256H1 (LEUKNOT03), 1974322H1 (UCMCL5T01), 2748718F6 (LUNGTUT11), 3472155H1 (LUNGNOT27), 568987H1 (MMLR3DT01), 6124407H1 (BRAHNON05), 6867412H1 (BRAGNON02) and 7979275H1 (LSUBDMC01) which are SEQ ID NOs:3-10, respectively.

[0053] The chemokine receptor-like protein comprising the amino acid sequence of SEQ ID NO:1, is 333 amino acids in length and has chemical and structural homology with human chemokine receptor (SEQ ID NO:12). **FIG. 2** shows the alignment between chemokine receptor-like protein and human chemokine receptor; the receptors share 26% identity. Both chemokine receptor-like protein and human chemokine receptor contain a G-protein receptor motif, $V_{108}$-$I_{124}$ and $A_{117}$-$I_{133}$, respectively. Designation as a GPCR is validated by PFAM, BLOCKS, PRINTS, all of which place the chemokine receptor protein in the rhodopsin-like GPCR superfamily. In addition, chemokine receptor-like protein and human chemokine receptor have potential amino terminal N-glycosylation sites at $N_{10}$, $N_{14}$, $N_{264}$, $N_{322}$, and $N_{329}$, potential phosphorylation sites at $S_{161}$, $S_{318}$, $S_{325}$, $T_{301}$, and $T_{313}$ and potential carboxy-terminal amidation sites at $M_{310}$ and $L_{312}$. They also share similar hydrophobicity plots as shown in U.S. Pat. No. 5,955,303, which is incorporated in its entirety by reference herein.

[0054] **FIGS. 3A and 3B** show the northern analysis for the cDNA encoding chemokine receptor-like protein. As can be seen in the last line of **FIG. 3**A, the chemokine receptor-like protein is rather sparsely expressed, 1.3×10e−8, and in its sparsity, closely resembles the expression pattern for other disease associated GPCRs in brain—normal expression during development and overexpressed in disorders such as cancer. Of particular note is the fact that in the nervous system, the percent abundance of the cDNA in brain cancers is approximately two-fold higher than expression in the brain tissue of the subject who died of CHF (chronic heart failure). Furthermore, sequence was never expressed in normal brain tissue or in brain tissues from subjects diagnosed with Alzheimer's (7 tissues), epilepsy (8 tissues), Huntington's chorea (16 tissues), or schizophrenia (9 tissues), or who died of CHF (27 tissues). Therefore, by expression pattern, the cDNA, the protein and antibody which specifically binds the protein are diagnostic of brain cancer, particularly meningioma.

[0055] The mRNA encoding chemokine receptor-like protein was also differentially expressed in peripheral blood mononuclear cell treated with Staphylococcal exotoxins. The experimental results which demonstrate the differential expression of this GPCR under conditions related to inflammation associated with bacterial infection are found in EXAMPLE VII.

[0056] A mammalian variant of the cDNA encoding chemokine receptor-like protein was identified using BLAST2 with default parameters and the ZOOSEQ databases (Incyte Genomics, Palo Alto Calif.). The rat variant, SEQ ID NO:11, has about 90% identity to the human sequence from about nucleotide 918 to about nucleotide 1229 of SEQ ID NO:2.

[0057] It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of cDNAs encoding chemokine receptor-like protein, some bearing minimal similarity to the cDNAs of any known and naturally occurring gene, may be produced. Thus, the inven-

tion contemplates each and every possible variation of cDNA that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide encoding naturally occurring chemokine receptor-like protein, and all such variations are to be considered as being specifically disclosed.

[0058] The cDNAs of SEQ ID NOs:2-10 may be used in hybridization, amplification, and screening technologies to identify and distinguish among SEQ ID NO:2 and related molecules in a sample. The mammalian cDNAs, SEQ ID NO:11, may be used to produce transgenic cell lines or organisms which are model systems for human disorders including infection, inflammation and cancer and upon which the toxicity and efficacy of potential therapeutic treatments may be tested. Toxicology studies, clinical trials, and subject/patient treatment profiles may be performed and monitored using the cDNAs, proteins, antibodies and molecules and compounds identified using the cDNAs and proteins of the present invention.

[0059] Characterization and Use of the Invention

[0060] cDNA Libraries

[0061] In a particular embodiment disclosed herein, mRNA is isolated from mammalian cells and tissues using methods which are well known to those skilled in the art and used to prepare the cDNA libraries. The Incyte cDNAs were isolated from mammalian cDNA libraries prepared as described in the EXAMPLES. The consensus sequences are chemically and/or electronically assembled from fragments including Incyte cDNAs and extension and/or shotgun sequences using computer programs such as PHRAP (Phil Green, University of Washington, Seattle Wash.), and the AUTOASSEMBLER application (Applied Biosystems (ABI), Foster City Calif.). After verification of the 5' and 3' sequence, at least one of the representative cDNAs which encode the chemokine receptor-like protein is designated a reagent. These reagent cDNAs are also used in the construction of human LIFEARRAYS (Incyte Genomics). The cDNA encoding the chemokine receptor-like protein was represented among the 17,719 sequences on LIFEGEM2 array (Incyte Genomics).

[0062] Sequencing

[0063] Methods for sequencing nucleic acids are well known in the art and may be used to practice any of the embodiments of the invention. These methods employ enzymes such as the Klenow fragment of DNA polymerase I, SEQUENASE, Taq DNA polymerase and thermostable T7 DNA polymerase (Amersham Pharmacia Biotech (APB), Piscataway N.J.), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad Calif.). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 system (Hamilton, Reno Nev.) and the DNA ENGINE thermal cycler (MJ Research, Watertown Mass.). Machines commonly used for sequencing include the PRISM 3700, 377 or 373 DNA sequencing systems (ABI), the MEGABACE 1000 DNA sequencing system (APB), and the like. The sequences may be analyzed using a variety of algorithms well known in the art and described in Ausubel et al. (1997; *Short Protocols in Molecular Biology*, John Wiley & Sons, New York N.Y., unit 7.7) and in Meyers (1995; *Molecular Biology and Biotechnology*, Wiley V C H, New York N.Y., pp. 856-853).

[0064] Shotgun sequencing may also be used to complete the sequence of a particular cloned insert of interest. Shotgun strategy involves randomly breaking the original insert into segments of various sizes and cloning these fragments into vectors. The fragments are sequenced and reassembled using overlapping ends until the entire sequence of the original insert is known. Shotgun sequencing methods are well known in the art and use thermostable DNA polymerases, heat-labile DNA polymerases, and primers chosen from representative regions flanking the cDNAs of interest. Incomplete assembled sequences are inspected for identity using various algorithms or programs such as CONSED (Gordon (1998) Genome Res 8:195-202) which are well known in the art. Contaminating sequences, including vector or chimeric sequences, or deleted sequences can be removed or restored, respectively, organizing the incomplete assembled sequences into finished sequences.

[0065] Extension of a Nucleic Acid Sequence

[0066] The sequences of the invention may be extended using various PCR-based methods known in the art. For example, the XL-PCR kit (ABI), nested primers, and commercially available cDNA or genomic DNA libraries may be used to extend the nucleic acid sequence. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO primer analysis software (Molecular Biology Insights, Cascade Colo.) to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to a target molecule at temperatures from about 55C. to about 68C. When extending a sequence to recover regulatory elements, it is preferable to use genomic, rather than cDNA libraries.

[0067] Hybridization

[0068] The cDNA and fragments thereof can be used in hybridization technologies for various purposes. A probe may be designed or derived from unique regions such as the 5' regulatory region or from a nonconserved region (i.e., 5' or 3' of the nucleotides encoding the conserved catalytic domain of the protein) and used in protocols to identify naturally occurring molecules encoding the chemokine receptor-like protein, allelic variants, or related molecules. The probe may be DNA or RNA, may be single-stranded, and should have at least 50% sequence identity to a nucleic acid sequence selected from SEQ ID NOs:2-11. Hybridization probes may be produced using oligolabeling, nick translation, end-labeling, or PCR amplification in the presence of a reporter molecule. A vector containing the cDNA or a fragment thereof may be used to produce an mRNA probe in vitro by addition of an RNA polymerase and labeled nucleotides. These procedures may be conducted using commercially available kits.

[0069] The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. Hybridization can be performed at low stringency with buffers, such as 5×SSC with 1% sodium dodecyl sulfate (SDS) at 60C., which permits the formation of a hybridization complex between nucleic acid sequences that contain some mismatches. Subsequent washes are performed at

higher stringency with buffers such as 0.2×SSC with 0.1% SDS at either 45C. (medium stringency) or 68C. (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acids are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide can be added to the hybridization solution to reduce the temperature at which hybridization is performed, and background signals can be reduced by the use of detergents such as Sarkosyl or TRITON X-100 (Sigma-Aldrich, St. Louis Mo.) and a blocking agent such as denatured salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (supra) and Sambrook et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview N.Y.

[**0070**] Arrays incorporating cDNAs may be prepared and analyzed using methods well known in the art. Oligonucleotides or cDNAs may be used as hybridization probes or targets to monitor the expression level of large numbers of genes simultaneously or to identify genetic variants, mutations, and single nucleotide polymorphisms. Monoclonal or polyclonal antibodies may be used to detect or quantify expression of a protein in a sample. Such arrays may be used to determine gene function; to understand the genetic basis of a condition, disease, or disorder; to diagnose a condition, disease, or disorder; and to develop and monitor the activities of therapeutic agents. (See, e.g., Brennan et al. (1995) U.S. Pat. No. 5,474,796; Schena et al. (1996) Proc Natl Acad Sci 93:10614-10619; Heller et al. (1997) Proc Natl Acad Sci 94:2150-2155; Heller et al. (1997) U.S. Pat. No. 5,605,662; and deWildt et al. (2000) Nature Biotechnol 18:989-994.)

[**0071**] Hybridization probes are also useful in mapping the naturally occurring genomic sequence. The probes may be hybridized to a particular chromosome, a specific region of a chromosome, or an artificial chromosome construction. Such constructions include human artificial chromosomes (HAC), yeast artificial chromosomes (YAC), bacterial artificial chromosomes (BAC), bacterial P1 constructions, or the cDNAs of libraries made from single chromosomes.

[**0072**] Quantitative PCR (TAQMAN, ABI)

[**0073**] Quantitative real-time PCR (QPCR) is a method for quantifying a nucleic acid molecule based on detection of a fluorescent signal produced during PCR amplification (Gibson et al. (1996) Genome Res 6:995-1001; Heid et al. (1996) Genome Res 6:986-994). Amplification is carried out on machines such as the ABI PRISM 7700 detection system which consists of a 96-well thermal cycler connected to a laser and charge-coupled device (CCD) optics system. To perform QPCR, a PCR reaction is carried out in the presence of a doubly labeled probe. The probe, which is designed to anneal between the standard forward and reverse PCR primers, is labeled at the 5' end by a flourogenic reporter dye such as 6-carboxyfluorescein (6-FAM) and at the 3' end by a quencher molecule such as 6-carboxy-tetramethylrhodamine (TAMRA). As long as the probe is intact, the 3' quencher extinguishes fluorescence by the 5' reporter. However, during each primer extension cycle, the annealed probe is degraded as a result of the intrinsic 5' to 3' nuclease activity of Taq polymerase (Holland et al. (1991) Proc Natl Acad Sci 88:7276-7280). This degradation separates the reporter from the quencher, and fluorescence is detected every few seconds by the CCD. The higher the starting copy number of the nucleic acid, the sooner a significant increase in fluorescence is observed. A cycle threshold ($C_T$) value, representing the cycle number at which the PCR product crosses a fixed threshold of detection is determined by the instrument software. The $C_T$ is inversely proportional to the copy number of the template and can therefore be used to calculate either the relative or absolute initial concentration of the nucleic acid molecule in the sample. The relative concentration of two different molecules can be calculated by determining their respective $C_T$ values (comparative $C_T$ method). Alternatively, the absolute concentration of the nucleic acid molecule can be calculated by constructing a standard curve using a housekeeping molecule of known concentration. The process of calculating $C_T$s, preparing a standard curve, and determining starting copy number is performed by the SEQUENCE DETECTOR 1.7 software (ABI).

[**0074**] Expression

[**0075**] Any one of a multitude of cDNAs encoding the chemokine receptor-like protein may be cloned into a vector and used to express the protein, or portions thereof, in host cells. The nucleic acid sequence can be engineered by such methods as DNA shuffling, as described in U.S. Pat. No. 5,830,721, and site-directed mutagenesis to create new restriction sites, alter glycosylation patterns, change codon preference to increase expression in a particular host, produce splice variants, extend half-life, and the like. The expression vector may contain transcriptional and translational control elements (promoters, enhancers, specific initiation signals, and polyadenylated 3' sequence) from various sources which have been selected for their efficiency in a particular host. The vector, cDNA, and regulatory elements are combined using in vitro recombinant DNA techniques, synthetic techniques, and/or in vivo genetic recombination techniques well known in the art and described in Sambrook (supra, ch. 4, 8, 16 and 17).

[**0076**] A variety of host systems may be transformed with an expression vector. These include, but are not limited to, bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems transformed with baculovirus expression vectors; plant cell systems transformed with expression vectors containing viral and/or bacterial elements, or animal cell systems (Ausubel supra, unit 16). For example, an adenovirus transcription/translation complex may be utilized in mammalian cells. After sequences are ligated into the E1 or E3 region of the viral genome, the infective virus is used to transform and express the protein in host cells. The Rous sarcoma virus enhancer or SV40 or EBV-based vectors may also be used for high-level protein expression.

[**0077**] Routine cloning, subcloning, and propagation of nucleic acid sequences can be achieved using the multifunctional pBLUESCRIPT vector (Stratagene, La Jolla Calif.) or pSPORT1 plasmid (Invitrogen). Introduction of a nucleic acid sequence into the multiple cloning site of these vectors disrupts the lacZ gene and allows calorimetric screening for transformed bacteria. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence.

[0078] For long term production of recombinant proteins, the vector can be stably transformed into cell lines along with a selectable or visible marker gene on the same or on a separate vector. After transformation, cells are allowed to grow for about 1 to 2 days in enriched media and then are transferred to selective media. Selectable markers, antimetabolite, antibiotic, or herbicide resistance genes, confer resistance to the relevant selective agent and allow growth and recovery of cells which successfully express the introduced sequences. Resistant clones identified either by survival on selective media or by the expression of visible markers may be propagated using culture techniques. Visible markers are also used to estimate the amount of protein expressed by the introduced genes. Verification that the host cell contains the desired cDNA is based on DNA-DNA or DNA-RNA hybridizations or PCR amplification techniques.

[0079] The host cell may be chosen for its ability to modify a recombinant protein in a desired fashion. Such modifications include acetylation, carboxylation, glycosylation, phosphorylation, lipidation, acylation and the like. Post-translational processing which cleaves a "prepro" form may also be used to specify protein targeting, folding, and/or activity. Different host cells available from the ATCC (Manassas Va.) which have specific cellular machinery and characteristic mechanisms for post-translational activities may be chosen to ensure the correct modification and processing of the recombinant protein.

[0080] Recovery of Proteins from Cell Culture

[0081] Heterologous moieties engineered into a vector for ease of purification include glutathione S-transferase (GST), 6×His, FLAG, MYC, and the like. GST and 6-His are purified using commercially available affinity matrices such as immobilized glutathione and metal-chelate resins, respectively. FLAG and MYC are purified using commercially available monoclonal and polyclonal antibodies. For ease of separation following purification, a sequence encoding a proteolytic cleavage site may be part of the vector located between the protein and the heterologous moiety. Methods for recombinant protein expression and purification are discussed in Ausubel (supra, unit 16) and are commercially available.

[0082] Protein Identification

[0083] Several techniques have been developed which permit rapid identification of proteins using high performance liquid chromatography and mass spectrometry. Beginning with a sample containing proteins, the major steps involved are: 1) proteins are separated using two-dimensional gel electrophoresis (2-DE), 2) selected proteins are excised from the gel and digested with a protease to produce a set of peptides; and 3) the peptides are subjected to mass spectral (MS) analysis to derive peptide ion mass and spectral pattern information. The MS information is used to identify the protein by comparing it with information in a protein database (Shevenko et al. (1996) Proc Natl Acad Sci 93:14440-14445).

[0084] Proteins are separated by 2DE employing isoelectric focusing (IEF) in the first dimension followed by SDS-PAGE in the second dimension. For IEF, an immobilzed pH gradient strip is useful to increase reproducibility and resolution of the separation. Alternative techniques may be used to improve resolution of very basic, hydrophobic, or high molecular weight proteins. The separated proteins are detected using a stain or dye such as silver stain, Coomassie blue, or spyro red (Molecular Probes, Eugene Oreg.) that is compatible with mass spectrometry Gels may be blotted onto a PVDF membrane for western analysis and optically scanned using a STORM scanner (APB) to produce a computer-readable output which is analyzed by pattern recognition software such as MELANIE (GeneBio, Geneva, Switzerland). The software annotates individual spots by assigning a unique identifier and calculating their respective x,y coordinates, molecular masses, isoelectric points, and signal intensity. Individual spots of interest, such as those representing differentially expressed proteins, are excised and proteolytically digested with a site-specific protease such as trypsin or chymotrypsin, singly or in combination, to generate a set of small peptides, preferably in the range of 1-2 kDa. Prior to digestion, samples may be treated with reducing and alkylating agents, and following digestion, the peptides are then separated by liquid chromatography or capillary electrophoresis and analyzed using MS.

[0085] MS converts components of a sample into gaseous ions, separates the ions based on their mass-to-charge ratio, and determines relative abundance. For peptide mass fingerprinting analysis, a mass spectrometer of the MALDI-TOF (Matrix Assisted Laser Desorption/Ionization-Time of Flight), ESI (Electrospray Ionization), and TOF-TOF (Time of Flight/Time of Flight) machines are used to determine a set of highly accurate peptide masses. Using analytical programs, such as TURBOSEQUEST software (Finnigan, San Jose Calif.), the MS data is compared against a database of theoretical MS data derived from known or predicted proteins. A minimum match of three peptide masses is usually required for reliable protein identification. If additional information is needed for identification, Tandem-MS may be used to derive information about individual peptides. In tandem-MS, a first stage of MS is performed to determine individual peptide masses. Then selected peptide ions are subjected to fragmentation using a technique such as collision induced dissociation (CID) to produce an ion series. The resulting fragmentation ions are analyzed in a second round of MS, and their spectral pattern may be used to determine a short stretch of amino acid sequence (Dancik et al. (1999) J Comput Biol 6:327-342).

[0086] Assuming the protein is represented in a searchable database, a combination of peptide mass and fragmentation data, together with the calculated MW and pI of the protein, will usually yield an unambiguous identification. If no match is found, protein sequence can be obtained using direct chemical sequencing procedures well known in the art (Creighton (1984) *Proteins, Structures and Molecular Properties*, W H Freeman, New York N.Y.).

[0087] Chemical Synthesis of Peptides

[0088] Proteins or portions thereof may be produced not only by recombinant methods, but also by using chemical methods well known in the art. Solid phase peptide synthesis may be carried out in a batchwise or continuous flow process which sequentially adds α-amino- and side chain-protected amino acid residues to an insoluble polymeric support via a linker group. A linker group such as methylamine-derivatized polyethylene glycol is attached to poly(styrene-co-divinylbenzene) to form the support resin. The amino acid residues are N-α-protected by acid labile Boc (t-butyloxy-

carbonyl) or base-labile Fmoc (9-fluorenylmethoxycarbonyl). The carboxyl group of the protected amino acid is coupled to the amine of the linker group to anchor the residue to the solid phase support resin. Trifluoroacetic acid or piperidine are used to remove the protecting group in the case of Boc or Fmoc, respectively. Each additional amino acid is added to the anchored residue using a coupling agent or pre-activated amino acid derivative, and the resin is washed. The full length peptide is synthesized by sequential deprotection, coupling of derivitized amino acids, and washing with dichloromethane and/or N, N-dimethylformamide. The peptide is cleaved between the peptide carboxy terminus and the linker group to yield a peptide acid or amide. (Novabiochem 1997/98 Catalog and Peptide Synthesis Handbook, San Diego Calif. pp. S1-S20). Automated synthesis may also be carried out on machines such as the 431A peptide synthesizer (ABI). A protein or portion thereof may be purified by preparative high performance liquid chromatography and its composition confirmed by amino acid analysis or by sequencing (Creighton (1984) *Proteins, Structures and Molecular Properties*, W H Freeman, New York N.Y.).

[**0089**] Antibodies

[**0090**] Antibodies, or immunoglobulins (Ig), are components of immune response expressed on the surface of or secreted into the circulation by B cells. The prototypical antibody is a tetramer composed of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds which binds and neutralizes foreign antigens. Based on their H-chain, antibodies are classified as IgA, IgD, IgE, IgG or IgM. The most common class, IgG, is tetrameric while other classes are variants or multimers of the basic structure.

[**0091**] Antibodies are described in terms of their two main functional domains. Antigen recognition is mediated by the Fab (antigen binding fragment) region of the antibody, while effector functions are mediated by the Fc (crystallizable fragment) region. The binding of antibody to antigen triggers destruction of the antigen by phagocytic white blood cells such as macrophages and neutrophils. These cells express surface Fc receptors that specifically bind to the Fc region of the antibody and allow the phagocytic cells to destroy antibody-bound antigen. Fc receptors are single-pass transmembrane glycoproteins containing about 350 amino acids whose extracellular portion typically contains two or three Ig domains (Sears et al. (1990) J Immunol 144:371-378).

[**0092**] Preparation and Screening of Antibodies

[**0093**] Various hosts including mice, rats, rabbits, goats, llamas, camels, and human cell lines may be immunized by injection with an antigenic determinant. Adjuvants such as Freund's, mineral gels, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemacyanin (KLH; Sigma-Aldrich), and dinitrophenol may be used to increase immunological response. In humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are preferable. The antigenic determinant may be an oligopeptide, peptide, or protein. When the amount of antigenic determinant allows immunization to be repeated, specific polyclonal antibody with high affinity can be obtained (Klinman and Press (1975) Transplant Rev 24:41-83). Oligopepetides which may contain between about five and about fifteen amino acids identical to a portion of the endogenous protein may be fused with proteins such as KLH in order to produce antibodies to the chimeric molecule.

[**0094**] Monoclonal antibodies may be prepared using any technique which provides for the production of antibodies by continuous cell lines in culture. These include the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler et al. (1975) Nature 256:495-497; Kozbor et al. (1985) J Immunol Methods 81:31-42; Cote et al. (1983) Proc Natl Acad Sci 80:2026-2030; and Cole et al. (1984) Mol Cell Biol 62:109-120).

[**0095**] Chimeric antibodies may be produced by techniques such as splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity (Morrison et al. (1984) Proc Natl Acad Sci 81:6851-6855; Neuberger et al. (1984) Nature 312:604-608; and Takeda et al. (1985) Nature 314:452-454). Alternatively, techniques described for antibody production may be adapted, using methods known in the art, to produce specific, single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton (1991) Proc Natl Acad Sci 88:10134-10137). Antibody fragments which contain specific binding sites for an antigenic determinant may also be produced. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse et al. (1989) Science 246:1275-1281).

[**0096**] Antibodies may also be produced by inducing production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in Orlandi et al. (1989; Proc Natl Acad Sci 86:3833-3837) or Winter et al. (1991; Nature 349:293-299). A protein may be used in screening assays of phagemid or B-lymphocyte immunoglobulin libraries to identify antibodies having a desired specificity. Numerous protocols for competitive binding or immunoassays using either polyclonal or monoclonal antibodies with established specificities are well known in the art.

[**0097**] Antibody Specificity

[**0098**] Various methods such as Scatchard analysis combined with radioimmunoassay techniques may be used to assess the affinity of particular antibodies for a protein. Affinity is expressed as an association constant, $K_a$, which is defined as the molar concentration of protein-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The $K_a$ determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple antigenic determinants, represents the average affinity, or avidity, of the antibodies. The $K_a$ determined for a preparation of monoclonal antibodies, which are specific for a particular antigenic determinant, represents a true measure of affinity. High-affinity antibody preparations with $K_a$ ranging from about $10^9$ to $10^{12}$ L/mole are preferred for use in immunoas-

says in which the protein-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with $K_a$ ranging from about $10^6$ to $10^7$ L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of the protein, preferably in active form, from the antibody (Catty (1988) *Antibodies, Volume I: A Practical Approach*, IRL Press, Washington D.C.; Liddell and Cryer (1991) *A Practical Guide to Monoclonal Antibodies*, John Wiley & Sons, New York N.Y.).

[0099] The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing about 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of protein-antibody complexes. Procedures for making antibodies, evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are widely available (Catty (supra); Ausubel (supra) pp. 11.1-11.31).

## DIAGNOSTICS

[0100] Disorders associated with differential expression of the chemokine-like receptor protein include infection, particularly complications of bacterial and viral infection; inflammation, particularly chronic ulcerative colitis, Crohn's disease, or complications of cancer; and cancers, particularly adenocarcinomas of the colon and prostate, brain tumors (meningioma, hypernephroma), breast tumors (ductal or intraductal), neuroganglion tumors (ganglioneuroma), small intestine tumors (carcinoid), transitional cell carcinoma of the bladder, and leiomyomata of the uterus. These conditions may be diagnosed or evaluated using any of the following assays.

[0101] Immunological Assays

[0102] Immunological methods for detecting and measuring complex formation as a measure of protein expression using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), fluorescence-activated cell sorting (FACS) and antibody arrays. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may be employed (Pound (1998) *Immunochemical Protocols*, Humana Press, Totowa N.J.). These assays and their quantitation against purified, labeled standards are well known in the art (Ausubel, supra, unit 10.1-10.6).

[0103] These methods are also useful for diagnosing diseases that show differential protein expression. Normal or standard values for protein expression are established by combining body fluids or cell extracts taken from a normal mammalian or human subject with specific antibodies to a protein under conditions for complex formation. Standard values for complex formation in normal and diseased tissues are established by various methods, often photometric means. Then complex formation as it is expressed in a subject sample is compared with the standard values. Deviation from the normal standard and toward the diseased standard provides parameters for disease diagnosis or prognosis while deviation away from the diseased and toward the normal standard may be used to evaluate treatment efficacy.

[0104] Recently, antibody arrays have allowed the development of techniques for high-throughput screening of recombinant antibodies. Such methods use robots to pick and grid bacteria containing antibody genes, and a filter-based ELISA to screen and identify clones that express antibody fragments. Because liquid handling is eliminated and the clones are arrayed from master stocks, the same antibodies can be spotted multiple times and screened against multiple antigens simultaneously. Antibody arrays are highly useful in the identification of differentially expressed proteins. (See de Wildt, supra)

[0105] Differential expression of chemokine receptor-like protein as detected using any of the above assays is diagnostic of infection, inflammation and cancer.

[0106] Nucleic Acid Assays

[0107] The cDNAs, fragments, oligonucleotides, complementary RNA and DNA molecules, and PNAs may be used to detect and quantify differential gene expression for diagnostic purposes. Similarly antibodies which specifically bind chemokine receptor-like protein may be used diagnostically, to quantitate protein expression. The diagnostic assay may use hybridization or amplification technology to compare gene expression in a biological sample from a patient to standard samples in order to detect differential gene expression. Qualitative or quantitative methods for this comparison are well known in the art.

[0108] Gene Expression Profiles

[0109] A gene expression profile comprises the expression of a plurality of cDNAs as measured by after hybridization with a sample. The cDNAs of the invention may be used as elements on a array to produce a gene expression profile. In one embodiment, the array is used to diagnose or monitor the progression of disease. Researchers can assess and catalog the differences in gene expression between healthy and diseased tissues or cells. For example, the cDNA or probe may be labeled by standard methods and added to a biological sample from a patient under conditions for the formation of hybridization complexes. After an incubation period, the sample is washed and the amount of label (or signal) associated with hybridization complexes, is quantified and compared with a standard value. If complex formation in the patient sample is significantly altered (higher or lower) in comparison to either a normal or disease standard, then differential expression indicates the presence of a disorder.

[0110] In order to provide standards for establishing differential expression, normal and disease expression profiles are established. This is accomplished by combining a sample taken from normal subjects, either animal or human, with a cDNA under conditions for hybridization to occur. Standard hybridization complexes may be quantified by comparing the values obtained using normal subjects with values from an experiment in which a known amount of a purified sequence is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who were diagnosed with a particular condition, disease, or disorder. Deviation from standard values toward those associated with a particular disorder is used to diagnose that disorder.

[0111] By analyzing changes in patterns of gene expression, disease can be diagnosed at earlier stages before the patient is symptomatic. The invention can be used to formulate a prognosis and to design a treatment regimen. The invention can also be used to monitor the efficacy of treatment. For treatments with known side effects, the array is employed to improve the treatment regimen. A dosage is established that causes a change in genetic expression patterns indicative of successful treatment. Expression patterns associated with the onset of undesirable side effects are avoided. This approach may be more sensitive and rapid than waiting for the patient to show inadequate improvement, or to manifest side effects, before altering the course of treatment.

[0112] In another embodiment, animal models which mimic a human disease can be used to characterize expression profiles associated with a particular condition, disease, or disorder; or treatment of the condition, disease, or disorder. Novel treatment regimens may be tested in these animal models using arrays to establish and then follow expression profiles over time. In addition, arrays may be used with cell cultures or tissues removed from animal models to rapidly screen large numbers of candidate drug molecules, looking for ones that produce an expression profile similar to those of known therapeutic drugs, with the expectation that molecules with the same expression profile will likely have similar therapeutic effects. Thus, the invention provides the means to rapidly determine the molecular mode of action of a drug.

[0113] Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies or in clinical trials or to monitor the treatment of an individual patient. Once the presence of a condition is established and a treatment protocol is initiated, diagnostic assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in a normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to years.

[0114] Labeling of Molecules for Assay

[0115] A wide variety of reporter molecules and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid, amino acid, and antibody assays. Synthesis of labeled molecules may be achieved using commercially available kits (Promega, Madison Wis.) for incorporation of a labeled nucleotide such as $^{32}$P-dCTP (APB), Cy3-dCTP or Cy5-dCTP (Qiagen-Operon, Alameda Calif.), or amino acid such as $^{35}$S-methionine (APB). Nucleotides and amino acids may be directly labeled with a variety of substances including fluorescent, chemiluminescent, or chromogenic agents, and the like, by chemical conjugation to amines, thiols and other groups present in the molecules using reagents such as BIODIPY or FITC (Molecular Probes, Eugene Oreg.).

## THERAPEUTICS

[0116] As described in THE INVENTION section, chemical and structural similarity, in particular sequence, specific motifs, or domains, exists between regions of the chemokine receptor-like protein (SEQ ID NO:1) and human chemokine receptor (SEQ ID NO:12) shown in **FIG. 2**. In addition, differential expression is highly associated with the inflammation due to bacterial infection as shown in EXAMPLE VII and brain cancer, particularly menigioma, as shown in **FIG. 3B**. The chemokine receptor-like protein clearly plays a role in these disorders.

[0117] In the treatment of cancer which is associated with the increased expression of the protein, it may be desirable to decrease protein expression or activity. In one embodiment, the an inhibitor, antagonist or antibody which specifically binds the protein may be administered to a subject to treat a condition associated with increased expression or activity. In another embodiment, a pharmaceutical composition comprising an inhibitor, antagonist, or antibody and a pharmaceutical carrier may be administered to a subject to treat a condition associated with the increased expression or activity of the endogenous protein. In an additional embodiment, a vector expressing the complement of the cDNA or fragments thereof may be administered to a subject to treat the disorder.

[0118] Any of the cDNAs, complementary molecules, or fragments thereof, proteins or portions thereof, vectors delivering these nucleic acid molecules or expressing the proteins, and their ligands may be administered in combination with other therapeutic agents. Selection of the agents for use in combination therapy may be made by one of ordinary skill in the art according to conventional pharmaceutical principles. A combination of therapeutic agents may act synergistically to affect treatment of a particular cancer at a lower dosage of each agent alone.

[0119] Modification of Gene Expression Using Nucleic Acids

[0120] Gene expression may be modified by designing complementary or antisense molecules (DNA, RNA, or PNA) to the control, 5', 3', or other regulatory regions of the gene encoding chemokine receptor-like protein. Oligonucleotides designed to inhibit transcription initiation are preferred. Similarly, inhibition can be achieved using triple helix base-pairing which inhibits the binding of polymerases, transcription factors, or regulatory molecules (Gee et al. In: Huber and Carr (1994) *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco N.Y., pp. 163-177). A complementary molecule may also be designed to block translation by preventing binding between ribosomes and mRNA. In one alternative, a library or plurality of cDNAs may be screened to identify those which specifically bind a regulatory, nontranslated sequence.

[0121] Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA followed by endonucleolytic cleavage at sites such as GUA, GUU, and GUC. Once such sites are identified, an oligonucleotide with the same sequence may be evaluated for secondary structural features which would render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing their hybridization with complementary oligonucleotides using ribonuclease protection assays.

[0122] Complementary nucleic acids and ribozymes of the invention may be prepared via recombinant expression, in vitro or in vivo, or using solid phase phosphoramidite chemical synthesis. In addition, RNA molecules may be

modified to increase intracellular stability and half-life by addition of flanking sequences at the 5' and/or 3' ends of the molecule or by the use of phosphorothioate or 2'O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Modification is inherent in the production of PNAs and can be extended to other nucleic acid molecules. Either the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, or the modification of adenine, cytidine, guanine, thymine, and uridine with acetyl-, methyl-, thio- groups renders the molecule less available to endogenous endonucleases.

[0123] cDNA Therapeutics

[0124] The cDNAs of the invention can be used in gene therapy. cDNAs can be delivered ex vivo to target cells, such as cells of bone marrow. Once stable integration and transcription and or translation are confirmed, the bone marrow may be reintroduced into the subject. Expression of the protein encoded by the cDNA may correct a disorder associated with mutation of a normal sequence, reduction or loss of an endogenous target protein, or overepression of an endogenous or mutant protein. Alternatively, cDNAs may be delivered in vivo using vectors such as retrovirus, adenovirus, adeno-associated virus, herpes simplex virus, and bacterial plasmids. Non-viral methods of gene delivery include cationic liposomes, polylysine conjugates, artificial viral envelopes, and direct injection of DNA (Anderson (1998) Nature 392:25-30; Dachs et al. (1997) Oncol Res 9:313-325; Chu et al. (1998) J Mol Med 76(3-4):184-192; Weiss et al. (1999) Cell Mol Life Sci 55(3):334-358; Agrawal (1996) *Antisense Therapeutics*, Humana Press, Totowa N.J.; and August et al. (1997) *Gene Therapy* (Advances in Pharmacology, Vol. 40), Academic Press, San Diego Calif.).

[0125] Screening and Purification Assays

[0126] The cDNA encoding chemokine receptor-like protein may be used to screen a library or a plurality of molecules or compounds for specific binding affinity. The libraries may be aptamers, DNA molecules, RNA molecules, PNAs, peptides, proteins such as transcription factors, enhancers, or repressors, and other ligands which regulate the activity, replication, transcription, or translation of the endogenous gene. The assay involves combining a polynucleotide with a library or plurality of molecules or compounds under conditions allowing specific binding, and detecting specific binding to identify at least one molecule which specifically binds the single-stranded or double-stranded molecule.

[0127] In one embodiment, the cDNA of the invention may be incubated with a plurality of purified molecules or compounds and binding activity determined by methods well known in the art, e.g., a gel-retardation assay (U.S. Pat. No. 6,010,849) or a reticulocyte lysate transcriptional assay. In another embodiment, the cDNA may be incubated with nuclear extracts from biopsied and/or cultured cells and tissues. Specific binding between the cDNA and a molecule or compound in the nuclear extract is initially determined by gel shift assay and may be later confirmed by recovering and raising antibodies against that molecule or compound. When these antibodies are added into the assay, they cause a supershift in the gel-retardation assay.

[0128] In another embodiment, the cDNA may be used to purify a molecule or compound using affinity chromatography methods well known in the art. In one embodiment, the cDNA is chemically reacted with cyanogen bromide groups on a polymeric resin or gel. Then a sample is passed over and reacts with or binds to the cDNA. The molecule or compound which is bound to the cDNA may be released from the cDNA by increasing the salt concentration of the flow-through medium and collected.

[0129] In a further embodiment, the protein or a portion thereof may be used to purify a ligand from a sample. A method for using a protein or a portion thereof to purify a ligand would involve combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and using a chaotropic agent to separate the protein from the purified ligand.

[0130] In a preferred embodiment, chemokine receptor-like protein may be used to screen a plurality of molecules or compounds in any of a variety of screening assays. The portion of the protein employed in such screening may be free in solution, affixed to an abiotic or biotic substrate (e.g. borne on a cell surface), or located intracellularly. For example, in one method, viable or fixed prokaryotic host cells that are stably transformed with recombinant nucleic acids that have expressed and positioned a peptide on their cell surface can be used in screening assays. The cells are screened against a plurality or libraries of ligands, and the specificity of binding or formation of complexes between the expressed protein and the ligand can be measured. Depending on the particular kind of molecules or compounds being screened, the assay may be used to identify DNA molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies, immunoglobulins, inhibitors, and drugs or any other ligand, which specifically binds the protein.

[0131] In one aspect, this invention comtemplates a method for high throughput screening using very small assay volumes and very small amounts of test compound as described in U.S. Pat. No. 5,876,946, incorporated herein by reference. This method is used to screen large numbers of molecules and compounds via specific binding. In another aspect, this invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound capable of binding to the protein. Molecules or compounds identified by screening may be used in a mammalian model system to evaluate their toxicity, diagnostic, or therapeutic potential.

[0132] Pharmaceutical Compositions

[0133] Pharmaceutical compositions may be formulated and administered, to a subject in need of such treatment, to attain a therapeutic effect. Such compositions contain the instant protein, agonists, antibodies specifically binding the protein, antagonists, inhibitors, or mimetics of the protein. Compositions may be manufactured by conventional means such as mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing. The composition may be provided as a salt, formed with acids such as hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic, or as a lyophilized powder which may be combined with a sterile buffer such as saline, dextrose, or water. These compositions may include auxiliaries or excipients which facilitate processing of the active compounds.

[0134] Auxiliaries and excipients may include coatings, fillers or binders including sugars such as lactose, sucrose, mannitol, glycerol, or sorbitol; starches from corn, wheat, rice, or potato; proteins such as albumin, gelatin and collagen; cellulose in the form of hydroxypropylmethyl-cellulose, methyl cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; lubricants such as magnesium stearate or talc; disintegrating or solubilizing agents such as the, agar, alginic acid, sodium alginate or cross-linked polyvinyl pyrrolidone; stabilizers such as carbopol gel, polyethylene glycol, or titanium dioxide; and dyestuffs or pigments added for identify the product or to characterize the quantity of active compound or dosage.

[0135] These compositions may be administered by any number of routes including oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal.

[0136] The route of administration and dosage will determine formulation; for example, oral administration may be accomplished using tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, or suspensions; parenteral administration may be formulated in aqueous, physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Suspensions for injection may be aqueous, containing viscous additives such as sodium carboxymethyl cellulose or dextran to increase the viscosity, or oily, containing lipophilic solvents such as sesame oil or synthetic fatty acid esters such as ethyl oleate or triglycerides, or liposomes. Penetrants well known in the art are used for topical or nasal administration.

[0137] Toxicity and Therapeutic Efficacy

[0138] A therapeutically effective dose refers to the amount of active ingredient which ameliorates symptoms or condition. For any compound, a therapeutically effective dose can be estimated from cell culture assays using normal and neoplastic cells or in animal models. Therapeutic efficacy, toxicity, concentration range, and route of administration may be determined by standard pharmaceutical procedures using experimental animals.

[0139] The therapeutic index is the dose ratio between therapeutic and toxic effects—LD50 (the dose lethal to 50% of the population)/ED50 (the dose therapeutically effective in 50% of the population)—and large therapeutic indices are preferred. Dosage is within a range of circulating concentrations, includes an ED50 with little or no toxicity, and varies depending upon the composition, method of delivery, sensitivity of the patient, and route of administration. Exact dosage will be determined by the practitioner in light of factors related to the subject in need of the treatment.

[0140] Dosage and administration are adjusted to provide active moiety that maintains therapeutic effect. Factors for adjustment include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular composition.

[0141] Normal dosage amounts may vary from 0.1 $\mu$g, up to a total dose of about 1 g, depending upon the route of administration. The dosage of a particular composition may be lower when administered to a patient in combination with other agents, drugs, or hormones. Guidance as to particular dosages and methods of delivery is provided in the pharmaceutical literature and generally available to practitioners. Further details on techniques for formulation and administration may be found in the latest edition of *Remington's Pharmaceutical Sciences* (Mack Publishing, Easton Pa.).

MODEL SYSTEMS

[0142] Animal models may be used as bioassays where they exhibit a phenotypic response similar to that of humans and where exposure conditions are relevant to human exposures. Mammals are the most common models, and most infectious agent, cancer, drug, and toxicity studies are performed on rodents such as rats or mice because of low cost, availability, lifespan, reproductive potential, and abundant reference literature. Inbred and outbred rodent strains provide a convenient model for investigation of the physiological consequences of under- or over-expression of genes of interest and for the development of methods for diagnosis and treatment of diseases. A mammal inbred to over-express a particular gene (for example, secreted in milk) may also serve as a convenient source of the protein expressed by that gene.

[0143] Toxicology

[0144] Toxicology is the study of the effects of agents on living systems. The majority of toxicity studies are performed on rats or mice. Observation of qualitative and quantitative changes in physiology, behavior, homeostatic processes, and lethality in the rats or mice are used to generate a toxicity profile and to assess potential consequences on human health following exposure to the agent.

[0145] Genetic toxicology identifies and analyzes the effect of an agent on the rate of endogenous, spontaneous, and induced genetic mutations. Genotoxic agents usually have common chemical or physical properties that facilitate interaction with nucleic acids and are most harmful when chromosomal aberrations are transmitted to progeny. Toxicological studies may identify agents that increase the frequency of structural or functional abnormalities in the tissues of the progeny if administered to either parent before conception, to the mother during pregnancy, or to the developing organism. Mice and rats are most frequently used in these tests because their short reproductive cycle allows the production of the numbers of organisms needed to satisfy statistical requirements.

[0146] Acute toxicity tests are based on a single administration of an agent to the subject to determine the symptomology or lethality of the agent. Three experiments are conducted: 1) an initial dose-range-finding experiment, 2) an experiment to narrow the range of effective doses, and 3) a final experiment for establishing the dose-response curve.

[0147] Subchronic toxicity tests are based on the repeated administration of an agent. Rat and dog are commonly used in these studies to provide data from species in different families. With the exception of carcinogenesis, there is considerable evidence that daily administration of an agent at high-dose concentrations for periods of three to four months will reveal most forms of toxicity in adult animals.

[0148] Chronic toxicity tests, with a duration of a year or more, are used to demonstrate either the absence of toxicity

or the carcinogenic potential of an agent. When studies are conducted on rats, a minimum of three test groups plus one control group are used, and animals are examined and monitored at the outset and at intervals throughout the experiment.

[0149] Transgenic Animal Models

[0150] Transgenic rodents that over-express or under-express a gene of interest may be inbred and used to model human diseases or to test therapeutic or toxic agents. (See, e.g., U.S. Pat. No. 5,175,383 and U.S. Pat. No. 5,767,337.) In some cases, the introduced gene may be activated at a specific time in a specific tissue type during fetal or postnatal development. Expression of the transgene is monitored by analysis of phenotype, of tissue-specific mRNA expression, or of serum and tissue protein levels in transgenic animals before, during, and after challenge with experimental drug therapies.

[0151] Embryonic Stem Cells

[0152] Embryonic (ES) stem cells isolated from rodent embryos retain the potential to form embryonic tissues. When ES cells are placed inside a carrier embryo, they resume normal development and contribute to tissues of the live-born animal. ES cells are the preferred cells used in the creation of experimental knockout and knockin rodent strains. Mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and are grown under culture conditions well known in the art. Vectors used to produce a transgenic strain contain a disease gene candidate and a marker gen, the latter serves to identify the presence of the introduced disease gene. The vector is transformed into ES cells by methods well known in the art, and transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains.

[0153] ES cells derived from human blastocysts may be manipulated in vitro to differentiate into at least eight separate cell lineages. These lineages are used to study the differentiation of various cell types and tissues in vitro, and they include endoderm, mesoderm, and ectodermal cell types which differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes.

[0154] Knockout Analysis

[0155] In gene knockout analysis, a region of a mammalian gene is enzymatically modified to include a non-mammalian gene such as the neomycin phosphotransferase gene (neo; Capecchi (1989) Science 244:1288-1292). The modified gene is transformed into cultured ES cells and integrates into the endogenous genome by homologous recombination. The inserted sequence disrupts transcription and translation of the endogenous gene. Transformed cells are injected into rodent blastulae, and the blastulae are implanted into pseudopregnant dams. Transgenic progeny are crossbred to obtain homozygous inbred lines which lack a functional copy of the mammalian gene. In one example, the mammalian gene is a human gene.

[0156] Knockin Analysis

[0157] ES cells can be used to create knockin humanized animals (pigs) or transgenic animal models (mice or rats) of human diseases. With knockin technology, a region of a human gene is injected into animal ES cells, and the human sequence integrates into the animal cell genome. Transformed cells are injected into blastulae and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of the analogous human condition. These methods have been used to model several human diseases.

[0158] Non-Human Primate Model

[0159] The field of animal testing deals with data and methodology from basic sciences such as physiology, genetics, chemistry, pharmacology and statistics. These data are paramount in evaluating the effects of therapeutic agents on non-human primates as they can be related to human health. Monkeys are used as human surrogates in vaccine and drug evaluations, and their responses are relevant to human exposures under similar conditions. Cynomolgus and Rhesus monkeys (*Macaca fascicularis* and *Macaca mulatta*, respectively) and Common Marmosets (*Callithrix jacchus*) are the most common non-human primates (NHPs) used in these investigations. Since great cost is associated with developing and maintaining a colony of NHPs, early research and toxicological studies are usually carried out in rodent models. In studies using behavioral measures such as drug addiction, NHPs are the first choice test animal. In addition, NHPs and individual humans exhibit differential sensitivities to many drugs and toxins and can be classified as a range of phenotypes from "extensive metabolizers" to "poor metabolizers" of these agents.

[0160] In additional embodiments, the cDNAs which encode the protein may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of cDNAs that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

## EXAMPLES

[0161] The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention. The preparation of the MMLR3DT01 cDNA library, in which Incyte clone 568987, was discovered is described.

[0162] I MMLR3DT01 cDNA Library Construction

[0163] The normal peripheral blood macrophages used for this library were obtained from two 24 year old, Caucasian males. This library represents a mixture of allogeneically stimulated human macrophage populations obtained from Ficoll/Hypaque purified buffy coats. The cells from the two different donors (not typed for HLA alleles) were incubated at a density of $1\times10^6$/ml for 72 hours in DME containing 10% human serum.

[0164] After incubation, the macrophages mostly adhered to the plastic surface of the petri dish, whereas most other cell types, B and T lymphocytes, remained in solution. The DME was decanted from the dish, and the dish was washed with phosphate buffered saline (PBS). Macrophages were released from the plastic surface by gently scraping the petri

dish in PBS/1 mM EDTA and lysed immediately in buffer containing guanidinium isothiocyanate.

[0165] The lysate was extracted twice with a mixture of acid phenol, pH 4.0, and centrifuged over a CsCl cushion using an SW28 rotor in a L8-70M ultracentrifuge (Beckman Coulter, Fullerton Calif.). The RNA was precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in water, and DNAse treated for 15 min at 37C. It must be noted that some contaminating T and B lymphocytes may have been present.

[0166] The RNA was used to make cDNAs using the SUPERSCRIPT plasmid system (Invitrogen) and the recommended protocol. The resulting cDNAs were fractionated on a SEPHAROSE CL4B column (APB), and those cDNAs exceeding 400 bp were ligated into the pSPORT plasmid (Invitrogen). The plasmid was transformed into chemically competent DH5α host cells (Invitrogen).

[0167] II Isolation and Sequencing of cDNA Clones

[0168] Plasmid DNA was released from the host cells and purified using the MINIPREP kit (Edge Biosytems, Gaithersburg Md.). The kit consists of a 96 well-block with reagents for 960 purifications. The recommended protocol was employed except for the following changes: 1) the 96 wells were each filled with only 1 ml of sterile TERRIFIC BROTH (BD Biosciences, Sparks Md.) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the bacteria were cultured for 24 hours and then lysed with 60 μl of lysis buffer; and 3) the block was centrifuged at 2900 rpm for 5 min in the GS-6R centrifuge (Beckman Coulter) before the contents of the block were added to the primary filter plate. An optional step of adding isopropanol to TRIS buffer was not routinely performed. After the last step in the protocol, samples were transferred to a 96-well block for storage.

[0169] The cDNAs were prepared for sequencing using the MICROLAB 2200 system (Hamilton) in combination with the DNA ENGINE thermal cyclers (MJ Research). The cDNAs were sequenced by the method of Sanger and Coulson (1975; J Mol Biol 94:441-448) using an PRISM 377 sequencing system (ABI) or the MEGABACE 1000 DNA sequencing system (APB). Most of the isolates were sequenced according to standard protocols and kits (ABI) with solution volumes of 0.25×-1.0× concentrations. In the alternative, cDNAs were sequenced using solutions and dyes from APB.

[0170] III Extension of cDNA Sequences

[0171] The cDNAs were extended using the cDNA clone and oligonucleotide primers. One primer was synthesized to initiate 5' extension of the known fragment, and the other, to initiate 3' extension of the known fragment. The initial primers were designed using commercially available primer analysis software to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68C. to about 72C. Any stretch of nucleotides that would result in hairpin structures and primer-primer dimerizations was avoided.

[0172] Selected cDNA libraries were used as templates to extend the sequence. If more than one extension was necessary, additional or nested sets of primers were designed. Preferred libraries have been size-selected to include larger cDNAs and random primed to contain more sequences with 5' or upstream regions of genes. Genomic libraries are used to obtain regulatory elements, especially extension into the 5' promoter binding region.

[0173] High fidelity amplification was obtained by PCR using methods such as that taught in U.S. Pat. No. 5,932,451. PCR was performed in 96-well plates using the DNA ENGINE thermal cycler (MJ Research). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing $Mg^{2+}$, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (APB), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B (Incyte Genomics): Step 1: 94C., three min; Step 2: 94C., 15 sec; Step 3: 60C., one min; Step 4: 68C., two min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68C., five min; Step 7: storage at 4C. In the alternative, the parameters for primer pair T7 and SK+ (Stratagene) were as follows: Step 1: 94C., three min; Step 2: 94C., 15 sec; Step 3: 57C., one min; Step 4: 68C., two min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68C., five min; Step 7: storage at 4C.

[0174] The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% reagent in 1×TE, v/v; Molecular Probes) and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Life Sciences, Acton Mass.) and allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1% agarose minigel to determine which reactions were successful in extending the sequence.

[0175] The extended clones were desalted, concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison Wis.), and sonicated or sheared prior to religation into pUC18 vector (APB). For shotgun sequences, the digested nucleotide sequences were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and the agar was digested with AGARACE enzyme (Promega). Extended clones were religated using T4 DNA ligase (New England Biolabs) into pUC18 vector (APB), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into *E. coli* competent cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37C. in 384-well plates in LB/2×carbenicillin liquid media.

[0176] The cells were lysed, and DNA was amplified using primers, Taq DNA polymerase (APB) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94C., three min; Step 2: 94C., 15 sec; Step 3: 60C., one min; Step 4: 72C., two min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72C., five min; Step 7: storage at 4C. DNA was quantified using PICOGREEN quantitation reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the conditions described above. Samples were diluted with 20% dimethylsulfoxide (DMSO; 1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the

DYENAMIC DIRECT cycle sequencing kit (APB) or the PRISM BIGDYE terminator cycle sequencing kit (ABI).

[0177] IV Homology Searching of cDNA Clones and Their Deduced Proteins

[0178] The cDNAs of the Sequence Listing or their deduced amino acid sequences were used to query databases such as GenBank, SwissProt, BLOCKS, and the like. These databases that contain previously identified and annotated sequences or domains were searched using BLAST or BLAST2 to produce alignments and to determine which sequences were exact matches or homologs. The alignments were to sequences of prokaryotic (bacterial) or eukaryotic (animal, fungal, or plant) origin. Alternatively, algorithms such as the one described in Smith and Smith (1992, Protein Engineering 5:35-51) could have been used to deal with primary sequence patterns and secondary structure gap penalties. All of the sequences disclosed in this application have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

[0179] As detailed in Karlin and Altschul (1993; Proc Natl Acad Sci 90:5873-5877), BLAST matches between a query sequence and a database sequence were evaluated statistically and only reported when they satisfied the threshold of $10^{-25}$ for nucleotides and $10^{-14}$ for peptides. Homology was also evaluated by product score calculated as follows: the % nucleotide or amino acid identity [between the query and reference sequences] in BLAST is multiplied by the % maximum possible BLAST score [based on the lengths of query and reference sequences] and then divided by 100. In comparison with hybridization procedures used in the laboratory, the stringency for an exact match was set from a lower limit of about 40 (with 1-2% error due to uncalled bases) to a 100% match of about 70.

[0180] The BLAST software suite (NCBI, Bethesda Md.), includes various sequence analysis programs including "blastn" that is used to align nucleotide sequences and BLAST2 that is used for direct pairwise comparison of either nucleotide or amino acid sequences. BLAST programs are commonly used with gap and other parameters set to default settings, e.g.: Matrix: BLOSUM62; Reward for match: 1; Penalty for mismatch: −2; Open Gap: 5 and Extension Gap: 2 penalties; Gap x drop-off: 50; Expect: 10; Word Size: 11; and Filter: on. Identity is measured over the entire length of a sequence. Brenner et al. (1998; Proc Natl Acad Sci 95:6073-6078, incorporated herein by reference) analyzed BLAST for its ability to identify structural homologs by sequence identity and found 30% identity is a reliable threshold for sequence alignments of at least 150 residues and 40%, for alignments of at least 70 residues.

[0181] The cDNAs of this application were compared with assembled consensus sequences or templates found in the LIFESEQ GOLD database (Incyte Genomics). Component sequences from cDNA, extension, full length, and shotgun sequencing projects were subjected to PHRED analysis and assigned a quality score. All sequences with an acceptable quality score were subjected to various pre-processing and editing pathways to remove low quality 3' ends, vector and linker sequences, polyA tails, Alu repeats, mitochondrial and ribosomal sequences, and bacterial contamination sequences. Edited sequences had to be at least 50 bp in length, and low-information sequences and repetitive elements such as dinucleotide repeats, Alu repeats, and the like, were replaced by "Ns" or masked.

[0182] Edited sequences were subjected to assembly procedures in which the sequences were assigned to gene bins. Each sequence could only belong to one bin, and sequences in each bin were assembled to produce a template. Newly sequenced components were added to existing bins using BLAST and CROSSMATCH. To be added to a bin, the component sequences had to have a BLAST quality score greater than or equal to 150 and an alignment of at least 82% local identity. The sequences in each bin were assembled using PHRAP. Bins with several overlapping component sequences were assembled using DEEP PHRAP. The orientation of each template was determined based on the number and orientation of its component sequences.

[0183] Bins were compared to one another, and those having local similarity of at least 82% were combined and reassembled. Bins having templates with less than 95% local identity were split. Templates were subjected to analysis by STITCHER/EXON MAPPER algorithms that determine the probabilities of the presence of splice variants, alternatively spliced exons, splice junctions, differential expression of alternative spliced genes across tissue types or disease states, and the like. Assembly procedures were repeated periodically, and templates were annotated using BLAST against GenBank databases such as GBpri. An exact match was defined as having from 95% local identity over 200 base pairs through 100% local identity over 100 base pairs and a homolog match as having an E-value (or probability score) of $\leqq 1\times10^{-8}$. The templates were also subjected to frameshift FASTx against GENPEPT, and homolog match was defined as having an E-value of $\leqq 1\times10^{-8}$. Template analysis and assembly was described in U.S. Ser. No. 09/276,534, filed Mar. 25, 1999.

[0184] Following assembly, templates were subjected to BLAST, motif, and other functional analyses and categorized in protein hierarchies using methods described in U.S. Ser. No. 08/812,290 and U.S. Ser. No. 08/811,758, both filed Mar. 6, 1997; in U.S. Ser. No. 08/947,845, filed Oct. 9, 1997; and in U.S. Ser. No. 09/034,807, filed Mar. 4, 1998. Then templates were analyzed by translating each template in all three forward reading frames and searching each translation against the PFAM database of hidden Markov model-based protein families and domains using the HMMER software package (Washington University School of Medicine, St. Louis Mo.). The cDNA was further analyzed using MAcDNASIS PRO software (Hitachi Software Engineering), and LASERGENE software (DNASTAR) and queried against public databases such as the GenBank rodent, mammalian, vertebrate, prokaryote, and eukaryote databases, SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

[0185] V Transcript Imaging

[0186] A transcript image was performed using the LIFESEQ GOLD database (September 2001 release, Incyte Genomics). This process allowed assessment of the relative abundance of the expressed polynucleotides in all of the cDNA libraries and was described in U.S. Pat. No. 5,840,484, incorporated herein by reference. All sequences and cDNA libraries in the LIFESEQ database were categorized by system, organ/tissue and cell type. The categories included cardiovascular system, connective tissue, digestive system, embryonic structures, endocrine system, exocrine

glands, female and male genitalia, germ cells, hemic/immune system, liver, musculoskeletal system, nervous system, pancreas, respiratory system, sense organs, skin, stomatognathic system, unclassified/mixed, and the urinary tract. Criteria for transcript imaging can be selected from category, number of cDNAs per library, library description, disease indication, clinical relevance of sample, and the like.

[0187] All sequences and cDNA libraries in the LIFESEQ database have been categorized by system, organ/tissue and cell type. For each category, the number of libraries in which the sequence was expressed were counted and shown over the total number of libraries in that category. For each library, the number of cDNAs were counted and shown over the total number of cDNAs in that library. In some transcript images, all normalized or subtracted libraries, which have high copy number sequences removed prior to processing, and all mixed or pooled tissues, which are considered non-specific in that they contain more than one tissue type or more than one subject's tissue, can be excluded from the analysis. Treated and untreated cell lines and/or fetal tissue data can also be excluded where clinical relevance is emphasized. Conversely, fetal tissue can be emphasized wherever elucidation of inherited disorders or differentiation of particular adult or embryonic stem cells into tissues or organs such as heart, kidney, nerves or pancreas would be aided by removing clinical samples from the analysis. Transcript imaging can also be used to support data from other methodologies such as guilt-by-association and hybridization analyses. The results of this analysis are presented in **FIGS. 3A and 3B**.

[0188] VI Chromosome Mapping

[0189] Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Genethon are used to determine if any of the cDNAs presented in the Sequence Listing have been mapped. Any of the fragments of the cDNA encoding chemokine receptor-like protein that have been mapped result in the assignment of all related regulatory and coding sequences to the same location. The genetic map locations are described as ranges, or intervals, of human chromosomes. The map position of an interval, in cM (which is roughly equivalent to 1 megabase of human DNA), is measured relative to the terminus of the chromosomal p-arm.

[0190] VII Hybridization Technologies and Analyses

[0191] Experimental Material and Design

[0192] Human peripheral blood mononuclear cells (PBMCs) can be classified into discrete cellular populations representing the major cellular components of the immune system. PBMCs contain about 52% lymphocytes (12% B lymphocytes; 40% T lymphocytes of 25% CD4+ and 15% CD8+; 20% NK cells; 25% monocytes; and 3% various cells that include dendritic cells and progenitor cells. The proportions and the biology of these cellular components vary slightly between healthy individuals depending on factors such as age, gender, past medical history, and genetic background.

[0193] Staphylococcus exotoxins (SEB) are secreted by the bacteria and specifically activate human T cells by expressing an appropriate TCR-Vβ chain. Although polyclonal in nature, T cells activated by SEB require antigen presenting cells (APCs) to present the exotoxin to the T cells and deliver the costimulatory signals required for optimum T cell activation. Although, SEB must be presented to T cells by APCs, these molecules are not always required to produce a response. In fact, SEBs bind directly to a non-polymorphic portion of the human MHC class II molecules, bypassing the need for capture, cleavage, and binding of the peptides to the polymorphic antigenic groove of the MHC class II molecules.

[0194] To evaluate differential gene expression, the PBMCs from a healthy donor was stimulated in vitro with SEB for 24 and 72 hours. The treated PBMCs were compared to PBMCs from the same donor cultured for 24 hours in the absence of SEB.

[0195] Selection of Sequences, Microarray Preparation and Use

[0196] In most cases, Incyte cDNAs represent template sequences derived from the LIFESEQ GOLD assembled human sequence database (Incyte Genomics). Where more than one clone was available for a particular template, the 5'-most clone in the template was used on the microarray. The cDNA encoding the chemokine receptor-like protein was among the 17,719 sequences on LIFEGEM2 array (Incyte Genomics).

[0197] To construct microarrays, cDNAs were amplified from bacterial cells using primers complementary to vector sequences flanking the cDNA insert. Thirty cycles of PCR increased the initial quantity of cDNAs from 1-2 ng to a final quantity greater than 5 μg. Amplified cDNAs were then purified using SEPHACRYL-400 columns (APB). Purified cDNAs were immobilized on polymer-coated glass slides. Glass microscope slides (Corning Life Sciences) were cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides were etched in 4% hydrofluoric acid (VWR Scientific Products, West Chester Pa.), washed thoroughly in distilled water, and coated with 0.05% aminopropyl silane (Sigma-Aldrich) in 95% ethanol. Coated slides were cured in a 110C. oven. cDNAs were applied to the coated glass substrate using a procedure described in U.S. Pat. No. 5,807,522. One microliter of the cDNA at an average concentration of 100 ng/μl was loaded into the open capillary printing element by a high-speed robotic apparatus which then deposited about 5 nl of cDNA per slide.

[0198] Microarrays were UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene), and then washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites were blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (Tropix, Bedford Mass.) for 30 minutes at 60C. followed by washes in 0.2% SDS and distilled water as before.

[0199] Preparation of Samples

[0200] The peripheral blood sample was obtained from Donor 3735, a 42 year old healthy Caucasian female.

[0201] The PBMC samples were lysed in 1 ml of TRIZOL reagent (Invitrogen). The lysates were vortexed thoroughly, incubated at room temperature for 2-3 minutes, and extracted with 0.5 ml chloroform. The extract was mixed,

incubated at room temperature for 5 minutes, and centrifuged at 15,000 rpm for 15 minutes at 4C. The aqueous layer was collected and an equal volume of isopropanol was added. Samples were mixed, incubated at room temperature for 10 minutes, and centrifuged at 15,000 rpm for 20 minutes at 4C. The supernatant was removed and the RNA pellet was washed with 1 ml of 70% ethanol, centrifuged at 15,000 rpm at 4C., and resuspended in RNAse-free water. The concentration of the total RNA was determined by measuring the optical density at 260 nm.

[0202] Poly(A) RNA was prepared using an OLIGOTEX mRNA kit (Qiagen, Valencia Calif.) with the following modifications: OLIGOTEX beads were washed in tubes instead of on spin columns, resuspended in elution buffer, and then loaded onto spin columns to recover mRNA. To obtain maximum yield, the mRNA was eluted twice.

[0203] Each poly(A) RNA sample was reverse transcribed using a cDNA synthesis system (Invitrogen) with Not I-T7-VN primers (5'GCATTAGC GGCCGCGAAATTAATAC-GACTCACTA TAGGGAGATTTTTTTT TTTTTTTTTTVN 3') and 100 units MMLV RNAseH (−) reverse-transcriptase (Progmega) in the first strand reaction. The resulting cDNA was purified on a CHROMASPIN TE-200 column (Clontech, Palo Alto Calif.) and lyophilized until dry. The cDNA was amplified 200-400 fold using an AMPLISCRIBE IVT kit (Epicentre Technologies, Madison Wis.) in a procedure modified from U.S. Pat. No. 5,716,785 and U.S. Pat. No. 5,891,636. The amplified RNA was purified on a CHROMASPIN DEPC-200 column (Clontech).

[0204] Amplified RNA was labeled using MMLV reverse-transcriptase, random primer (9 mer), 1× first strand buffer, 0.03 units/μl RNAse inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, and 40 μM either dCTP-Cy3 or dCTP-Cy5 (APB). The reverse transcription reaction was performed in a 25 ml volume containing 200 ng poly(A) RNA using the GEMBRIGHT kit (Incyte Genomics). Specific control poly(A) RNAs (YCFR06, YCFR45, YCFR67, YCFR85, YCFR43, YCFR22, YCFR23, YCFR25, YCFR44, YCFR26) were synthesized by in vitro transcription from non-coding yeast genomic DNA (W. Lei, unpublished). As quantitative controls, control mRNAs (YCFR06, YCFR45, YCFR67, and YCFR85) at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng were diluted into reverse transcription reaction at ratios of 1:100,000, 1:10,000, 1:1000, 1:100 (w/w) to sample mRNA, respectively. To sample differential expression patterns, control mRNAs (YCFR43, YCFR22, YCFR23, YCFR25, YCFR44, YCFR26) were diluted into reverse transcription reaction at ratios of 1:3, 3:1, 1:10, 10:1, 1:25, 25:1 (w/w) to sample mRNA. Reactions were incubated at 37° C. for 2 hr, treated with 2.5 ml of 0.5M sodium hydroxide, and incubated for 20 minutes at 85° C. to the stop the reaction and degrade the RNA.

[0205] cDNAs were purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech). Cy3- and Cy5-labeled reaction samples were combined as described below and ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The cDNAs were then dried to completion using a SPEEDVAC system (Savant Instruments, Holbrook N.Y.) and resuspended in 14 μl 5×SSC/0.2% SDS.

[0206] Hybridization and Detection

[0207] Competitive hybridization reactions compared cDNAs derived from treated and untreated PMBCs from the same donor. cDNAs prepared from the samples were hybridized to the LIFEGEM2.

[0208] Hybridization reactions contained 9 μl of sample mixture containing 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5×SSC, 0.2% SDS hybridization buffer. The mixture was heated to 65C. for 5 minutes and was aliquoted onto the microarray surface and covered with an 1.8 $cm^2$ coverslip. The microarrays were transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber was kept at 100% humidity internally by the addition of 140 μl of 5×SSC in a corner of the chamber. The chamber containing the microarrays was incubated for about 6.5 hours at 60C. The microarrays were washed for 10 min at 45° C. in low stringency wash buffer (1×SSC, 0.1% SDS), three times for 10 minutes each at 45° C. in high stringency wash buffer (0.1×SSC), and dried.

[0209] Reporter-labeled hybridization complexes were detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Santa Clara Calif.) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light was focused on the microarray using a 20× microscope objective (Nikon, Melville N.Y.). The slide containing the microarray was placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm×1.8 cm microarray used in the present example was scanned with a resolution of 20 micrometers.

[0210] In two separate scans, the mixed gas multiline laser excited the two fluorophores sequentially. Emitted light was split, based on wavelength, into two photomultiplier tube detectors (PMT R1477; Hamamatsu Photonics Systems, Bridgewater N.J.) corresponding to the two fluorophores. Appropriate filters positioned between the microarray and the photomultiplier tubes were used to filter the signals. The emission maxima of the fluorophores used were 565 nm for Cy3 and 650 nm for CyS. Each microarray was typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus was capable of recording the spectra from both fluorophores simultaneously.

[0211] The sensitivity of the scans was calibrated using the signal intensity generated by a cDNA control species. Samples of the calibrating cDNA were separately labeled with the two fluorophores and identical amounts of each were added to the hybridization mixture. A specific location on the microarray contained a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000.

[0212] The output of the photomultiplier tube was digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Norwood, Mass.) installed in an IBM-compatible PC computer.

[0213] The digitized data were displayed as an image where the signal intensity was mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data was also analyzed quantitatively. Where two different fluorophores

were excited and measured simultaneously, the data were first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

[0214] A grid was superimposed over the fluorescence signal image such that the signal from each spot was centered in each element of the grid. The fluorescence signal within each element was then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis was the GEMTOOLS gene expression analysis program (Incyte Genomics). Significance was defined as signal to background ratio exceeding 2× and area hybridization exceeding 40%.

[0215] Experimental Results

[0216] The results of microarray experiments comparing untreated and SEB treated PBMCs from healthy Donor 3735 are:

| log2(Cy5/Cy3) | Sample (Cy3) | Sample (Cy5) |
|---|---|---|
| 1.053251 | PBMC Cells, Untx, | PBMC Cells, t/SEB 1 ng/ml 24 hr |

[0217] These results show a differential and significant increase in the expression of the mRNA encoding chemokine receptor-like protein after treatment pof PBMCs with SEB, an inflammatory response that would follow infection and production of exotoxin.

[0218] VIII Complementary Molecules

[0219] Molecules complementary to the cDNA, from about 5 (PNA) to about 5000 bp (complement of a cDNA insert), are used to detect or inhibit gene expression. Detection is described in Example VII. To inhibit transcription by preventing promoter binding, the complementary molecule is designed to bind to the most unique 5' sequence and includes nucleotides of the 5' UTR upstream of the initiation codon of the open reading frame. Complementary molecules include genomic sequences (such as enhancers or introns) and are used in "triple helix" base pairing to compromise the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. To inhibit translation, a complementary molecule is designed to prevent ribosomal binding to the mRNA encoding the protein.

[0220] Complementary molecules are placed in expression vectors and used to transform a cell line to test efficacy; into an organ, tumor, synovial cavity, or the vascular system for transient or short term therapy; or into a stem cell, zygote, or other reproducing lineage for long term or stable gene therapy. Transient expression lasts for a month or more with a non-replicating vector and for three months or more if elements for inducing vector replication are used in the transformation/expression system.

[0221] Stable transformation of dividing cells with a vector encoding the complementary molecule produces a transgenic cell line, tissue, or organism (U.S. Pat. No. 4,736,866). Those cells that assimilate and replicate sufficient quantities of the vector to allow stable integration also produce enough complementary molecules to compromise or entirely eliminate activity of the cDNA encoding the protein.

[0222] IX Expression of Chemokine Receptor-Like Protein

[0223] Expression and purification of the protein are achieved using either a mammalian cell expression system or an insect cell expression system. The pUB6/V5-His vector system (Invitrogen) is used to express chemokine receptor-like protein in CHO cells. The vector contains the selectable bsd gene, multiple cloning sites, the promoter/enhancer sequence from the human ubiquitin C gene, a C-terminal V5 epitope for antibody detection with anti-V5 antibodies, and a C-terminal polyhistidine (6×His) sequence for rapid purification on PROBOND resin (Invitrogen). Transformed cells are selected on media containing blasticidin.

[0224] *Spodoptera frugiperda* (Sf9) insect cells are infected with recombinant *Autographica californica* nuclear polyhedrosis virus (baculovirus). The polyhedrin gene is replaced with the cDNA by homologous recombination and the polyhedrin promoter drives cDNA transcription. The protein is synthesized as a fusion protein with 6×his which enables purification as described above. Purified protein is used in the following activity and to make antibodies

[0225] X Production of Antibodies

[0226] Chemokine receptor-like protein is purified using polyacrylamide gel electrophoresis and used to immunize mice or rabbits. Antibodies are produced using the protocols well known in the art and summarized below. Alternatively, the amino acid sequence of chemokine receptor-like protein is analyzed using LASERGENE software (DNASTAR) to determine regions of high antigenicity. An antigenic epitope, usually found near the C-terminus or in a hydrophilic region is selected, synthesized, and used to raise antibodies. Typically, epitopes of about 15 residues in length are produced using an 431A peptide synthesizer (ABI) using Fmoc-chemistry and coupled to KLH (Sigma-Aldrich) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester to increase antigenicity. Rabbits are immunized with the epitope-KLH complex in complete Freund's adjuvant.

[0227] Immunizations are repeated at intervals thereafter in incomplete Freund's adjuvant. After a minimum of seven weeks for mouse or twelve weeks for rabbit, antisera are drawn and tested for antipeptide activity. Testing involves binding the peptide to plastic, blocking with 1% bovine serum albumin, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG. Methods well known in the art are used to determine antibody titer and the amount of complex formation.

[0228] XI Immunopurification Using Specific Antibodies

[0229] Naturally occurring or recombinant protein is purified by immunoaffinity chromatography using antibodies which specifically bind the protein. An immunoaffinity column is constructed by covalently coupling the antibody to CNBr-activated SEPHAROSE resin (APB). Media containing the protein is passed over the immunoaffinity column, and the column is washed using high ionic strength buffers in the presence of detergent to allow preferential absorbance of the protein. After coupling, the protein is eluted from the column using a buffer of pH 2-3 or a high concentration of urea or thiocyanate ion to disrupt antibody/protein binding, and the protein is collected.

[0230] XII Antibody Arrays

[0231] Protein:protein Interactions

[0232] In an alternative to yeast two hybrid system analysis of proteins, an antibody array can be used to study protein-protein interactions and phosphorylation. A variety of protein ligands are immobilized on a membrane using methods well known in the art. The array is incubated in the presence of cell lysate until protein:antibody complexes are formed. Proteins of interest are identified by exposing the membrane to an antibody specific to the protein of interest. In the alternative, a protein of interest is labeled with digoxigenin (DIG) and exposed to the membrane; then the membrane is exposed to anti-DIG antibody which reveals where the protein of interest forms a complex. The identity of the proteins with which the protein of interest interacts is determined by the position of the protein of interest on the membrane.

[0233] Proteomic Profiles

[0234] Antibody arrays can also be used for high-throughput screening of recombinant antibodies. Bacteria containing antibody genes are robotically-picked and gridded at high density (up to 18,342 different double-spotted clones) on a filter. Up to 15 antigens at a time are used to screen for clones to identify those that express binding antibody fragments. These antibody arrays can also be used to identify proteins which are differentially expressed in samples (de Wildt, supra)

[0235] XIII Screening Molecules for Specific Binding with the cDNA or Protein

[0236] The cDNA, or fragments thereof, or the protein, or portions thereof, are labeled with $^{32}$P-dCTP, Cy3-dCTP, or Cy5-dCTP (APB), or with BIODIPY or FITC (Molecular Probes, Eugene Oreg.), respectively. Libraries of candidate molecules or compounds previously arranged on a substrate are incubated in the presence of labeled cDNA or protein. After incubation under conditions for either a nucleic acid or amino acid sequence, the substrate is washed, and any position on the substrate retaining label, which indicates specific binding or complex formation, is assayed, and the ligand is identified. Data obtained using different concentrations of the nucleic acid or protein are used to calculate affinity between the labeled nucleic acid or protein and the bound molecule.

[0237] XIV Two-Hybrid Screen

[0238] A yeast two-hybrid system, MATCHMAKER LexA Two-Hybrid system (Clontech), is used to screen for peptides that bind the protein of the invention. A cDNA encoding the protein is inserted into the multiple cloning site of a pLexA vector, ligated, and transformed into *E. coli*. cDNA, prepared from mRNA, is inserted into the multiple cloning site of a pB42AD vector, ligated, and transformed into *E. coli* to construct a cDNA library. The pLexA plasmid and pB42AD-cDNA library constructs are isolated from *E. coli* and used in a 2:1 ratio to co-transform competent yeast EGY48[p8op-lacZ] cells using a polyethylene glycol/lithium acetate protocol. Transformed yeast cells are plated on synthetic dropout (SD) media lacking histidine (−His), tryptophan (−Trp), and uracil (−Ura), and incubated at 30C. until the colonies have grown up and are counted. The colonies are pooled in a minimal volume of 1×TE (pH 7.5), replated on SD/−His/−Leu/−Trp/−Ura media supplemented with 2% galactose (Gal), 1% raffinose (Raf), and 80 mg/ml 5-bromo-4-chloro-3-indolyl β-d-galactopyranoside (X-Gal), and subsequently examined for growth of blue colonies. Interaction between expressed protein and cDNA fusion proteins activates expression of a LEU2 reporter gene in EGY48 and produces colony growth on media lacking leucine (−Leu). Interaction also activates expression of β-galactosidase from the p8op-lacZ reporter construct that produces blue color in colonies grown on X-Gal.

[0239] Positive interactions between expressed protein and cDNA fusion proteins are verified by isolating individual positive colonies and growing them in SD/−Trp/−Ura liquid medium for 1 to 2 days at 30C. A sample of the culture is plated on SD/−Trp/−Ura media and incubated at 30C. until colonies appear. The sample is replica-plated on SD/−Trp/−Ura and SD/−His/−Trp/−Ura plates. Colonies that grow on SD containing histidine but not on media lacking histidine have lost the pLexA plasmid. Histidine-requiring colonies are grown on SD/Gal/Raf/X-Gal/−Trp/−Ura, and white colonies are isolated and propagated. The pB42AD-cDNA plasmid, which contains a cDNA encoding a protein that physically interacts with the protein, is isolated from the yeast cells and characterized.

[0240] XV Demonstration of Chemokine Receptor-Like Protein Activity

[0241] GTP-binding activity is assayed by incubating varying amounts of chemokine receptor-like protein for 10 minutes at 30C. in 50 mM Tris buffer, pH 7.5, containing 1 mM dithiothreitol, 1 mM EDTA, 1 $\mu$M (a-$^{32}$P), in the absence or presence of 100 $\mu$M of the following compounds: GTP, GDP, GTPγS, ATP, CTP, UTP, and TTP. Samples are passed through nitrocellulose filters and washed twice with a buffer containing 50 mM Tris-HCL, pH 7.8, 1 mM $NaN_3$, 10 mM $MgCl_2$, 1 mM EDTA, 0.5 mM dithiothreitol, 0.01 mM PMSF, and 200 mM NaCl. The filter-bound counts are determined by liquid scintillation.

[0242] To determine GTPase activity, chemokine receptor-like protein is incubated for 10 minutes at 37C. in 50 mM Tris-HCL buffer, pH 7.8, containing 1 mM dithiothreitol, 2 mM EDTA, 10 $\mu$M (a-$^{32}$P), and 1 $\mu$M H-rab protein. GTPase activity is initiated by adding $MgCl_2$ to a final concentration of 10 mM. Samples are removed at various time points, mixed with an equal volume of ice-cold 0.5 mM EDTA, and frozen. Aliquots are spotted onto polyethyleneimine-cellulose thin layer chromatography plates, which are developed in 1M LiCl, dried, and autoradiographed.

[0243] All patents and publications mentioned in the specification are incorporated by reference herein. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 12

<210> SEQ ID NO 1
<211> LENGTH: 333
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 568987CD1

<400> SEQUENCE: 1

```
Met Asn Thr Thr Val Met Gln Gly Phe Asn Arg Ser Glu Arg Cys
  1               5                  10                  15

Pro Arg Asp Thr Arg Ile Val Gln Leu Val Phe Pro Ala Leu Tyr
                 20                  25                  30

Thr Val Val Phe Leu Thr Gly Ile Leu Leu Asn Thr Leu Ala Leu
                 35                  40                  45

Trp Val Phe Val His Ile Pro Ser Ser Ser Thr Phe Ile Ile Tyr
                 50                  55                  60

Leu Lys Asn Thr Leu Val Ala Asp Leu Ile Met Thr Leu Met Leu
                 65                  70                  75

Pro Phe Lys Ile Leu Ser Asp Ser His Leu Ala Pro Trp Gln Leu
                 80                  85                  90

Arg Ala Phe Val Cys Arg Phe Ser Ser Val Ile Phe Tyr Glu Thr
                 95                 100                 105

Met Tyr Val Gly Ile Val Leu Leu Gly Leu Ile Ala Phe Asp Arg
                110                 115                 120

Phe Leu Lys Ile Ile Arg Pro Leu Arg Asn Ile Phe Leu Lys Lys
                125                 130                 135

Pro Val Phe Ala Lys Thr Val Ser Ile Phe Ile Trp Phe Phe Leu
                140                 145                 150

Phe Phe Ile Ser Leu Pro Ile Met Ile Leu Ser Asn Lys Glu Ala
                155                 160                 165

Thr Pro Ser Ser Val Lys Lys Cys Ala Ser Leu Lys Gly Pro Leu
                170                 175                 180

Gly Leu Lys Trp His Gln Met Val Asn Asn Ile Cys Gln Phe Ile
                185                 190                 195

Phe Trp Thr Val Leu Ile Leu Met Leu Val Phe Tyr Val Val Ile
                200                 205                 210

Ala Lys Lys Val Tyr Asp Ser Tyr Arg Lys Ser Lys Cys Lys Asp
                215                 220                 225

Arg Lys Asn Asn Lys Lys Leu Glu Gly Lys Val Phe Val Val Val
                230                 235                 240

Pro Val Phe Phe Val Cys Phe Ala Pro Phe His Phe Ala Arg Val
                245                 250                 255

Pro Tyr Thr His Ser Gln Thr Asn Asn Lys Thr Asp Cys Arg Leu
                260                 265                 270

Gln Asn Gln Leu Phe Ile Ala Lys Glu Thr Thr Leu Phe Leu Ala
                275                 280                 285

Ala Thr Asn Ile Cys Met Asp Pro Leu Ile Ser Ile Phe Leu Cys
                290                 295                 300

Lys Lys Phe Thr Glu Lys Leu Pro Cys Met Gln Gly Arg Lys Thr
```

-continued

305 310 315

Thr Ala Ser Ser Gln Glu Asn His Ser Ser Gln Thr Asp Asn Ile
320 325 330

Thr Leu Gly

<210> SEQ ID NO 2
<211> LENGTH: 1488
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 568987CB1

<400> SEQUENCE: 2

actagttcaa gaggccatct acgaacgtat gactgccgct ttaagaagac agagagaact 60

gagtatcctc ccaaaggtga cactggaagc aatgaacacc acagtgatgc aaggcttcaa 120

cagatctgag cggtgcccca gagacactcg gatagtacag ctggtattcc cagccctcta 180

cacagtggtt ttcttgaccg gcatcctgct gaatactttg gctctgtggg tgtttgttca 240

catccccagc tcctccacct tcatcatcta cctcaaaaac actttggtgg ccgacttgat 300

aatgacactc atgcttcctt tcaaaatcct ctctgactca cacctggcac cctggcagct 360

cagagctttt gtgtgtcgtt tttcttcggt gatattttat gagaccatgt atgtgggcat 420

cgtgctgtta gggctcatag cctttgacag attcctcaag atcatcagac ctttgagaaa 480

tatttttcta aaaaaacctg tttttgcaaa aacggtctca atcttcatct ggttcttttt 540

gttcttcatc tccctgccaa ttatgatctt gagcaacaag gaagcaacac catcgtctgt 600

gaaaaagtgt gcttccttaa aggggcctct ggggctgaaa tggcatcaaa tggtaaataa 660

catatgccag tttattttct ggactgtttt aatcctaatg cttgtgtttt atgtggttat 720

tgcaaaaaaa gtatatgatt cttatagaaa gtccaaatgt aaggacagaa aaaacaacaa 780

aaagctggaa ggcaaagtat ttgttgtcgt gcctgtcttc tttgtgtgtt ttgctccatt 840

tcattttgcc agagttccat atactcacag tcaaaccaac aataagactg actgtagact 900

gcaaaatcaa ctgtttattg ctaaagaaac aactctcttt ttggcagcaa ctaacatttg 960

tatggatccc ttaatatcca tattcttatg taaaaaattc acagaaaagc taccatgtat 1020

gcaagggaga aagaccacag catcaagcca agaaaatcat agcagtcaga cagacaacat 1080

aaccttaggc tgacaactgt acatagggtt aacttctatt tattgatgag acttccgtag 1140

ataatgtgga aatcaaattt aaccaagaaa aaaagattgg aacaaatgct ctcttacatt 1200

ttattatcct ggtgtacaga aaagattata taaaatttaa atccacatag atctattcat 1260

aagctgaatg aaccattact aagagaatgc aacaggatac aaatggccac tagaggtcat 1320

tatttctttc tttctttttt ttttttttta atttcaagag catttcactt taacattttg 1380

gaaaagacta aggagaaacg tatatcccta caaacctccc ctctaaacac cttctcacat 1440

ttttttccac aattcacata acactactgc ttttgtcccc ttaaatgt 1488

<210> SEQ ID NO 3
<211> LENGTH: 257
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1881256H1

-continued

<400> SEQUENCE: 3

attcttttcc acaattcaca taacactact gcttttgtgc cccttaaatg tagatatgtg 60

ctgaaagaaa aaaaaaacgc ccaactcttg aagtccattg ctgaaaactg cagccagggg 120

ttgaaaggga tgcagacttg aagagtctga ggaactgaag tgggtcagca agacctctga 180

aatcctgggt aaaggatttt ctccttacaa ttacaaacag cctctttcac attacaataa 240

tataccatag gaggcac 257

<210> SEQ ID NO 4
<211> LENGTH: 268
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 1974322H1
<221> NAME/KEY: unsure
<222> LOCATION: 100-137
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 4

agaaaagatt atataaaatt taaatccaca tagatctatt cataagctga atgaaccatt 60

actaagagaa tgcaacagga tacaaatggc cactagaggn nnnnnnnnnn nnnnnnnnnn 120

nnnnnnnnnn nnnnnnncaa gagcatttca ctttaacatt ttggaaaaga ctaaggagaa 180

acgtatatcc ctacaaacct cccctccaaa caccttctca cattcttttc cacaattcac 240

ataacactac tgcttttgtg ccccttaa 268

<210> SEQ ID NO 5
<211> LENGTH: 359
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 2748718F6
<221> NAME/KEY: unsure
<222> LOCATION: 20, 26, 129, 144, 200, 298, 333, 344
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 5

aagcaacacc atcgtctgtn aaaaantgtg cttccttaaa ggggcctctg gggctgaaat 60

ggcatcaaat ggtaaataac atatgccagt ttattttctg gactgttttt atcctaatgc 120

ttgtgtttna tgtggttatt gcanaaaaag tatatggatt cttatagaaa gtccaaaagt 180

aaggacagaa aaaacaacan aaagctggaa ggcaaagtat ttgttgtcgt ggctgtcttc 240

tttgtgtgtt ttgctccatt tcattttgcc agagttccat atactcacag tcaaaccnac 300

aataagactg gctgtagact gcaaaatcaa ctngttattg ctanagaaac aactctctt 359

<210> SEQ ID NO 6
<211> LENGTH: 243
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 3472155H1

<400> SEQUENCE: 6

taatatacat attcttatgt aaaaaattca cagaaaagct accatgtatg caagggagaa 60

-continued

agaccacagc atcaagccaa gaaaatcata gcagtcagac agacaacata accttaggct 120

gacaactgta catagggtta acttctattt attgatgaga cttccgtaga taatgtggaa 180

atcaaattta accaagaaaa aaagattgga acaaatgctc tcttacattt tattatcctg 240

gtg 243

<210> SEQ ID NO 7
<211> LENGTH: 250
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 568987H1

<400> SEQUENCE: 7

aaaaaaagta tatgattctt atagaaagtc caaaagtaag gacagaaaaa acaacaaaaa 60

gctggaaggc aaagtatttg ttgtcgtggc tgtcttcttt gtgtgttttg ctccatttca 120

ttttgccaga gttccatata ctcacagtca aaccaacaat aagactgact gtagactgca 180

aaatcaactg tttattgcta aagaaacaac tctcttttttg gcagcaacta acatttgtat 240

ggatccctta 250

<210> SEQ ID NO 8
<211> LENGTH: 182
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 6124407H1
<221> NAME/KEY: unsure
<222> LOCATION: 123-124
<223> OTHER INFORMATION: a, t, c, g, or other

<400> SEQUENCE: 8

caatttaacc aagaaaaaaa gattggaaca aatgctctct tacattttat tatcctggtg 60

tacagaaaag attatattaa atttaaattc cacatagatc tattcattaa gctgaatgaa 120

ccnnattact aagagaatgc gaacaggata ccaaatggcc cactagaagg tccattattt 180

ct 182

<210> SEQ ID NO 9
<211> LENGTH: 592
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 6867412H1

<400> SEQUENCE: 9

gcttcaacag atctaagcgg tgccccagag acactcggat agtacagctg gtattcccag 60

ccctctacac agtggttttc ttgaccggca tcctgctgaa tactttggct ctgtgggtgt 120

ttgttcacat ccccagctcc tccaccttca tcatctacct caaaaacact ttggtggccg 180

acttgataat gacactcatg cttcctttca aaatcctctc tgactcacac ctggcaccct 240

ggcagctcag agcttttgtg tgtcgttttt cttcggtgat attttatgag accatgtatg 300

tgggcatcgt gctgttaggg ctcatagcct ttgacagatt cctcaagatc atcagacctt 360

tgagaaatat ttttctaaaa aaacctgttt ttgcaaaaac ggtctcaatc ttcatctggt 420

-continued

tctttttgtt cttcatctcc ctgccaaata tgatcttgag caacaaggaa gcaacaccat 480

cgtctgtgaa aaagtgtgct tccttaaagg ggcctctggg gctgaaatgg catcaaatgg 540

taaataacat atgccagttt attttctgga ctgttttatc ctaatgcttg tg 592

<210> SEQ ID NO 10
<211> LENGTH: 518
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 7979275H1

<400> SEQUENCE: 10

gggggctcat ttgtaggctg aactaatgac tgccgccata agaagacaga gagaactgag 60

tatcctccca aaggtgacac tggaagcaat gaacaccaca gtgatgcaag gcttcaacag 120

atctgagcgg tgccccagag acactcggat agtacagctg gtattcccag ccctctacac 180

agtggttttc ttgaccggca tcctgctgaa tactttggct ctgtgggtgt ttgttcacat 240

ccccagctcc tccaccttca tcatctacct caaaaacact ttggtggccg acttgataat 300

gacactcatg cttcctttca aaatcctctc tgactcacac ctggcaccct ggcagctcag 360

agcttttgtg tgtcgttttt cttcggtgat attttatgag accatgtatg tgggcatcgt 420

gctgttaggg ctcatagcct ttgacagatt cctcaagatc atcagacctt tgagaaatat 480

ttttctaaaa aaacctgttt ttgcaaaaac ggtctcaa 518

<210> SEQ ID NO 11
<211> LENGTH: 667
<212> TYPE: DNA
<213> ORGANISM: Rattus norvegicus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: 326157_Rn.1

<400> SEQUENCE: 11

gactgtagat tagaaaacca gctgtgtctt gctaaagaat caactctctt cctggcaaca 60

actaacattt gtatggaccc cttaatatat atcatcttgt gtaagaagtt cacccggaag 120

gtaccatgta tgagatggag gacaaagaca gcggcgtcca gcgatgagca ccacagcagt 180

cagacagaca acatcaccct atcctgacca ctttgtccca caggctaatt tcacacattt 240

ttctatgtga ggataggtct tcaaaaggcc atttacgtgg agacttcatt taagcattac 300

aggaaaaaaa agaggggaac aaacagtttc ctacatttta ttatcctcgt gtacggaaaa 360

gattatgccc attttaacca catagctgta tttgcaagca ggatgaatta acattaagag 420

aacatgtaat aaagcaaatg accactagat gtcacctttt caagaacatt cgtgtaatta 480

tggaaacagt taatgggaaa caggtttgcc taaaaaaaac ctcccttcta gttaccatcc 540

catgttctca cacacacaca agtccaaaaa catcatgctg ggtttttata gcctttagaa 600

tgcagacact tacggacaga aaccaacaga cttgtatatc cagtgcctgt acaggaaagg 660

gtggggg 667

<210> SEQ ID NO 12
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:

-continued

<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Incyte ID No: g992700

<400> SEQUENCE: 12

```
Met Asn Gly Leu Glu Val Ala Pro Pro Gly Leu Ile Thr Asn Phe
  1               5                  10                  15

Ser Leu Ala Thr Ala Glu Gln Cys Gly Gln Glu Thr Pro Leu Glu
                 20                  25                  30

Asn Met Leu Phe Ala Ser Phe Tyr Leu Leu Asp Phe Ile Leu Ala
                 35                  40                  45

Leu Val Gly Asn Thr Leu Ala Leu Trp Leu Phe Ile Arg Asp His
                 50                  55                  60

Lys Ser Gly Thr Pro Ala Asn Val Phe Leu Met His Leu Ala Val
                 65                  70                  75

Ala Asp Leu Ser Cys Val Leu Val Leu Pro Thr Arg Leu Val Tyr
                 80                  85                  90

His Phe Ser Gly Asn His Trp Pro Phe Gly Glu Ile Ala Cys Arg
                 95                 100                 105

Leu Thr Gly Phe Leu Phe Tyr Leu Asn Met Tyr Ala Ser Ile Tyr
                110                 115                 120

Phe Leu Thr Cys Ile Ser Ala Asp Arg Phe Leu Ala Ile Val His
                125                 130                 135

Pro Val Lys Ser Leu Lys Leu Arg Arg Pro Leu Tyr Ala His Leu
                140                 145                 150

Ala Cys Ala Phe Leu Trp Val Val Val Ala Val Ala Met Ala Pro
                155                 160                 165

Leu Leu Val Ser Pro Gln Thr Val Gln Thr Asn His Thr Val Val
                170                 175                 180

Cys Leu Gln Leu Tyr Arg Glu Lys Ala Ser His His Ala Leu Val
                185                 190                 195

Ser Leu Ala Val Ala Phe Thr Phe Pro Phe Ile Thr Thr Val Thr
                200                 205                 210

Cys Tyr Leu Leu Ile Ile Arg Ser Leu Arg Gln Gly Leu Arg Val
                215                 220                 225

Glu Lys Arg Leu Lys Thr Lys Ala Val Arg Met Ile Ala Ile Val
                230                 235                 240

Leu Ala Ile Phe Leu Val Cys Phe Val Pro Tyr His Val Asn Arg
                245                 250                 255

Ser Val Tyr Val Leu His Tyr Arg Ser His Gly Ala Ser Cys Ala
                260                 265                 270

Thr Gln Arg Ile Leu Ala Leu Ala Asn Arg Ile Thr Ser Cys Leu
                275                 280                 285

Thr Ser Leu Asn Gly Ala Leu Asp Pro Ile Met Tyr Phe Phe Val
                290                 295                 300

Ala Glu Lys Phe Arg His Ala Leu Cys Asn Leu Leu Cys Gly Lys
                305                 310                 315

Arg Leu Lys Gly Pro Pro Pro Ser Phe Glu Gly Lys Thr Asn Glu
                320                 325                 330

Ser Ser Leu Ser Ala Lys Ser Glu Leu
                335
```

What is claimed is:

**1**. An isolated cDNA comprising a nucleic acid sequence encoding a protein having the amino acid sequence of SEQ ID NO:1.

**2**. A purified protein comprising an amino acid sequence of SEQ ID NO:1.

**3**. A biologically active portion of the protein of claim 2 consisting of residues V108 -I133.

**4**. A composition comprising the protein of claim 2.

**5**. A method for using a protein to screen a plurality of compounds to identify at least one ligand, the method comprising:

a) combining the protein of claim 2 with the compounds under conditions to allow specific binding; and

b) detecting specific binding, thereby identifying a ligand that specifically binds the protein.

**6**. The method of claim 5 wherein the compounds are selected from agonists, antagonists, DNA molecules, mimetics, peptides, peptide nucleic acids, proteins, and RNA molecules.

**7**. A method of using a protein to identify an antibody that specifically binds the protein comprising:

a) contacting a plurality of antibodies with the protein of claim 2 under conditions to allow specific binding,

b) detecting specific binding between an antibody and the protein thereby identifying an antibody that specifically binds the protein.

**8**. A purified antibody that specifically binds the cehmokine receptor-like protein produced by the method of claim 7.

**9**. The antibody of claim 7, wherein the antibody is selected from an intact immunoglobulin molecule, a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a recombinant antibody, a humanized antibody, a single chain antibody, a Fab fragment, an $F(ab')_2$ fragment, an Fv fragment; and an antibody-peptide fusion protein.

**10**. A method of using a protein to prepare and purify a polyclonal antibody comprising:

a) immunizing a animal with a protein of claim 2 under conditions to elicit an antibody response;

b) isolating animal antibodies;

c) attaching the protein to a substrate;

d) contacting the substrate with isolated antibodies under conditions to allow specific binding to the protein;

e) dissociating the antibodies from the protein, thereby obtaining purified polyclonal antibodies.

**11**. A polyclonal antibody produced by the method of claim 10.

**12**. A method of using a protein to prepare a monoclonal antibody comprising:

a) immunizing a animal with a protein of claim 2 under conditions to elicit an antibody response;

b) isolating antibody-producing cells from the animal;

c) fusing the antibody-producing cells with immortalized cells in culture to form monoclonal antibody producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating monoclonal antibodies from culture.

**13**. A monoclonal antibody produced by the method of claim 12.

**14**. A method for using an antibody to immunopurify a protein comprising:

a) attaching the antibody of claim 8 to a substrate,

b) exposing the antibody to a sample containing protein under conditions to allow antibody:protein complexes to form,

c) dissociating the protein from the complex, and

d) collecting the purified protein.

**15**. A method for using an antibody to detect expression of a protein in a sample, the method comprising:

a) combining the antibody of claim 8 with a sample under conditions which allow the formation of antibody:protein complexes; and

b) detecting complex formation, wherein complex formation indicates expression of the protein in the sample.

**16**. The method of claim 15 wherein the sample is brain tissue or peripheral blood.

**17**. The method of claim 15 wherein complex formation is compared with standards and is diagnostic of meningioma.

**18**. The method of claim 15 wherein complex formation is compared with standards and is diagnostic of Staphylococcus infection.

**19**. A composition comprising an antibody of claim 8 and a labeling moiety or a pharmaceutical agent.

**20**. A antagonist which specifically binds the protein of claim 2.

* * * * *

| | | | |
|---|---|---|---|
| 专利名称(译) | GPCR诊断脑癌 | | |
| 公开(公告)号 | US20030165989A1 | 公开(公告)日 | 2003-09-04 |
| 申请号 | US10/100982 | 申请日 | 2002-03-18 |
| [标]申请(专利权)人(译) | AU YOUNG JANICE<br>GUEGLER KARLĴ<br>程MUZONG | | |
| 申请(专利权)人(译) | AU-YOUNG JANICE<br>GUEGLER KARL J.<br>程MUZONG | | |
| 当前申请(专利权)人(译) | AU-YOUNG JANICE<br>GUEGLER KARL J.<br>程MUZONG | | |
| [标]发明人 | AU YOUNG JANICE<br>GUEGLER KARL J<br>CHENG MUZONG | | |
| 发明人 | AU-YOUNG, JANICE<br>GUEGLER, KARL J.<br>CHENG, MUZONG | | |
| IPC分类号 | A61K38/00 C07K14/715 G01N33/53 C07H21/04 C12P21/04 C12P21/02 C12N5/06 | | |
| CPC分类号 | C07K14/7158 A61K38/00 | | |
| 外部链接 | Espacenet USPTO | | |


摘要(译)

本发明提供趋化因子受体样蛋白，编码该蛋白的cDNA和特异性结合该蛋白的抗体。它还提供cDNA，蛋白质和抗体在感染，炎症和癌症，特别是脑膜脑瘤的诊断，预后，治疗和评估中的用途。本发明还提供了用于产生蛋白质和转基因模型系统的载体和宿主细胞。


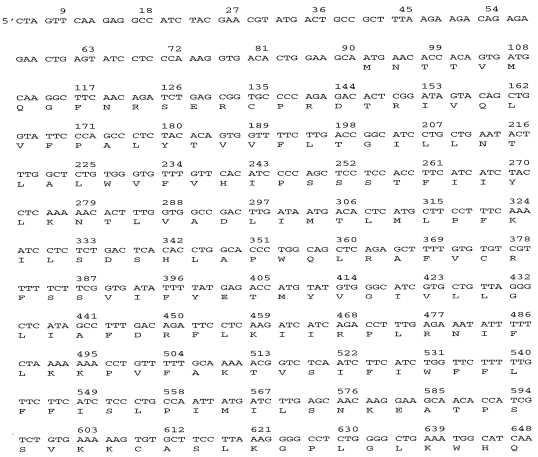

FIGURE 1A