

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 February 2002 (07.02.2002)

PCT

(10) International Publication Number
WO 02/10392 A2

(51) International Patent Classification⁷: **C12N 15/12**,
5/12, 5/10, C07K 14/705, 16/28, G01N 33/50, A61K 38/17

(21) International Application Number: PCT/US01/24351

(22) International Filing Date: 1 August 2001 (01.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/222,624 1 August 2000 (01.08.2000) US

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier application:
US 60/222,624 (CON)
Filed on 1 August 2000 (01.08.2000)

(71) Applicants (for all designated States except US):
OREGON HEALTH & SCIENCE UNIVERSITY
[US/US]; 3181 S.W. Sam Jackson Park Road, Port-
land, OR 97201-3098 (US). **THE GOVERNMENT**
OF THE UNITED STATES doing business as THE
DEPARTMENT OF VETERANS AFFAIRS [US/US];
Department of Veteran Affairs, 810 Vermont Avenue,
Washington, DC 20420 (US). **TARGETED GENE DE-**
LIVERY [US/US]; Suite 200, 2611 SW 3rd Avenue,
Portland, OR 97201 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HEFENEIDER,**
Steven [US/US]; 7220 SW Northvale Way, Portland, OR

97225 (US). **MERKINS, Louise, S.** [US/US]; 3181 S.W.
Sam Jackson Park Road, Portland, OR 97201 (US). **BEN-**
NETT, Robert, M. [US/US]; 3181 S.W. Sam Jackson
Park Road, Portland, OR 97201 (US). **SEISS, Donald, C.**
[US/US]; 3181 S.W. Sam Jackson Park Road, Portland,
OR 97201 (US).

(74) Agent: **NOONAN, Kevin, E.**; McDonnell Boehnen Hul-
bert & Berghoff, 300 South Wacker Drive, Chicago, IL
60606 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: MAMMALIAN DNA BINDING MEMBRANE-ASSOCIATED PROTEIN-ENCODING GENE AND USES

(57) Abstract: The present invention relates to novel mammalian DNA-R proteins and genes that encode such proteins. The invention is directed toward the isolation and characterization of mammalian DNA-R proteins. The invention specifically provides isolated complementary DNA copies of mRNA corresponding to rat and human homologues of a mammalian DNA-R gene. Also provided are recombinant expression constructs capable of expressing the mammalian DNA-R genes of the invention in cultures of transformed prokaryotic and eukaryotic cells, as well as such cultures of transformed cells that synthesise the mammalian catecholamine receptor proteins encoded therein. The invention also provides methods for screening compounds (*in vitro*) that are capable of binding to the mammalian DNA-R proteins of the invention, and further characterizing the binding properties of such compounds in comparison with known DNA-R agonists and antagonists. Improved methods of pharmacological screening are provided thereby.



WO 02/10392 A2

**MAMMALIAN DNA BINDING MEMBRANE-ASSOCIATED
PROTEIN-ENCODING GENE AND USES**

This application claims priority to U.S. Provisional Application Serial No.
5 60/222,624, filed August 1, 2000, the entirety of which is incorporated by reference.

BACKGROUND OF THE INVENTION

10 **1. Field of the Invention**

This invention relates to cell membrane-associated DNA binding proteins (termed DNA-R herein) from mammalian species and the genes corresponding to such receptors. Specifically, the invention relates to the isolation, cloning and sequencing of complementary DNA (cDNA) copies of messenger RNA (mRNA) encoding a novel mammalian DNA-R gene. The invention
15 also relates to the construction of recombinant expression constructs comprising cDNA of this novel DNA-R gene, said recombinant expression constructs being capable of expressing DNA-R protein in cultures of transformed prokaryotic and eukaryotic cells. Production of the receptor protein in such cultures is also provided, as well as the production of fragments thereof having biological activity. The invention relates to the use of such cultures of such transformed cells
20 to produce homogeneous compositions of the novel DNA-R protein. The invention also provides cultures of such cells producing this DNA-R protein for the characterization of novel and useful drugs. Antibodies against and epitopes of this novel DNA-R protein are also provided by the invention.

25 **2. Background of the Invention**

Extracellular DNA is a potent biological signal, being capable of initiating a wide range of immune responses *in vivo* and *in vitro*; including cytokine production, influx of neutrophils, IgM secretion, B-cell proliferation and enhanced natural killer activity. These properties of extracellular DNA enable naked DNA to be used as vaccines, in some instances. In addition,
30 extracellular DNA has been used to introduce new genetic information into cells, both *in vivo* and *in vitro*.

One important aspect of extracellular DNA transfer into mammalian cells is gene therapy. Gene transfer therapy offers the potential for treatment of a variety of diseases. The ability to provide safe, efficient, and selective *in vivo* gene delivery will be a critical component
35 of future protocols. Gene transfer by injection of either plasmid DNA or DNA/liposome

complexes has been demonstrated to be safe and permits expression of gene products. The uptake of DNA/liposome complexes does not depend upon specific cell-surface receptors while the mechanism mediating uptake of plasmid DNA by cells remains unknown.

5 In order to realize the full potential of this technology, safe delivery and efficient transgene expression of DNA in selected tissues and cells must be achieved. One approach to target DNA to tissue is the use of a receptor-mediated mechanism for the binding and internalization of DNA. Viral (retrovirus, adenovirus, adeno-associated virus) delivery of DNA to cells is via a receptor-mediated mechanism, however this technique has limited *in vivo* clinical application. Viral vectors have been most frequently used for *ex vivo* gene therapy, but 10 the technical problems associated with transplanting transduced cells remain a serious obstacle. In addition, viral vectors have the potential to lead to virus infection or to induce an immune response against antigenic viral coat proteins.

Non-viral methods of gene delivery include liposomes, the so-called "gene gun", and direct injection. Gene transfer with liposomes has been shown to result in uptake and 15 expression of DNA. Although DNA/liposomes are effectively taken up and the cDNA on the plasmid expressed, the process is believed to be nonspecific with limited possibility of targeting selected tissue. An alternative is to administer plasmid DNA directly, without a delivery system. Cells lines in tissue culture have demonstrated *in vitro* uptake of plasmid DNA and the expression of the transgene on the plasmid. It has also been shown that DNA, injected directly 20 *in vivo*, has been taken up and the encoded genes have been expressed. While this approach has been shown to be a safe and free from problems associated with DNA delivery by viruses, the therapeutic potential of this technology is often limited by poor transgene expression from plasmid DNA in many tissues. In addition, the mechanism by which plasmid DNA is bound and internalized into cells is not well established. Knowledge of the mechanism of plasmid 25 DNA binding to the cell surface, and how DNA is internalized and expressed, will be critical to enhancing transgene methods that also have the potential to target selected tissues.

Antisense oligonucleotides (ODN) are another form of extracellular DNA of great importance. ODN are considered potential therapeutic agents against various pathogens and oncogenes due to their ability to specifically inhibit gene expression. When injected into 30 tissues, ODN are internalized by cells and bind to complementary region of mRNA to inhibit translation of proteins in a highly specific manner. Different antisense ODN to HIV RNA have been shown to inhibit the infectivity of the virus in cultured human leukemia cells. Although human clinical trials using ODN to treat AIDS and other diseases are ongoing, the lack of a

precise understanding of where and how gene expression is effected hinders the optimization of this technique.

Extracellular DNA is also associated with human diseases, such as cystic fibrosis. Cystic fibrosis (CF) is the most common lethal genetic disease in North America. It affects one in 2500 live births and affected individuals have a median life expectancy of 28 years (Davis *et al.*, 1996, *Amer. J. Respir. Crit Care Med.* 157: 1234-1239). There is a growing body of evidence showing that inflammation, particularly the injurious products of neutrophils, may be responsible for lung damage (Doring, 1997, *Ped. Pulmonol. Supp.* 16: 271-272); it is now recognized that most of the morbidity and over 90% of the mortality results from chronic progressive inflammation of the lungs. Corticosteroids have a broad anti-inflammatory effect, particularly on neutrophils. A multicenter trial showed beneficial effects of oral corticosteroids on lung function. However, adverse effects such as growth retardation, glucose abnormalities and cataracts prelude this treatment as a long-term option (Eigen *et al.*, 1995, *J. Ped.* 126: 515-523). The nonsteroidal anti-inflammatory drug, ibuprofen, has also been studied (Konstan *et al.*, 1995, *N. Engl. J. Med.* 332: 848-854). The drug is beneficial, but continued monitoring is needed to determine the safety of long-term, high dose therapy. Other therapies that treat the injurious products of neutrophils, for example, antiproteases and antioxidants, are currently under investigation (Konstan, 1998, *Clin. Chest Med.* 19: 505-513).

The vicious airway fluid characteristic of CF can obstruct airflow and provides a viable growth medium for pathogenic bacteria, and cell lysis of these bacteria can produce extracellular DNA that causes inflammation. Recombinant human Dnase (rhDNase) has been clinical use since 1994 (Kontsan, 1998, *ibid.*). The rhDNase, administered by inhalation, has been used to cleave the extracelular airway DNA and reduce the viscosity of the airway fluid. Treatment with rhDNase produces a small improvement in lung function (Cramer & Bosso, 1996, *Ann. Pharmacol.* 30: 656-661). However, when treatment is stopped, patients can deteriorate to a point below their previous baseline (Bush, 1998, *Ped. Pulmonol.* 25: 79-82). In addition, a recent report showed that despite improvements in lung function, there were no changes in airway inflammation (Henry *et al.*, 1998, *Ped. Pulmonol.* 26: 97-100). Although the DNA is broken down by the Dnase, it is not entirely degraded, and hydrolized fragments are still potentially immunostimulatory and can contribute to inflammation. Thus rhDNase may be masking the process of on-going lung destruction.

There are also a variety of conventional treatments for CF including physiotherapy, nutritional support and drugs (Bilton & Mahadeva, 1997, *J. Royal Soc. Med.* 90: Suppl.31, 2-5).

Because the events that trigger and sustain inflammation in patients with CF are not clearly understood, a variety of approaches have been developed to treat different components of the disease. Antibiotics, anti-inflammatories, and therapies to reduce the viscosity of the airway fluid are all approaches that are being used and investigated. Aggressive antibiotic therapy has helped the acute control of infection, but rarely if ever are the bacteria in the airways of patients with CF completely eradicated. These pathogenic bacteria chronically stimulate and exacerbate inflammation. Although some of the currently-available treatments can help to alleviate symptoms and slow the progression of disease, none of the current treatments can prevent ultimate respiratory failure.

One important clinical observation is that greatly increased amounts of extracellular DNA, of host and bacterial origin, are present in the airway of patients with cystic fibrosis. Recent investigation has demonstrated that extracellular DNA, purified from sputum of patients with CF, will directly induce inflammation in the mouse lung (Schwartz *et al.*, 1997, *J. Clin. Invest.* 100: 68-73). The DNA purified from the sputum of patients with cystic fibrosis has been shown to be composed primarily of host-derived DNA and only a small fraction appears to be bacterial DNA (Schwartz *et al.*, 1997, *ibid.*). One possible explanation is that extracellular DNA binds to immune lung cells in the lungs and induces the secretion of pro-inflammatory cytokines and neutrophil migration to the lung, leading to severe airway inflammation. Extracellular DNA binding to immune cells in the lung, such as alveolar macrophages are stimulated to produce pro-inflammatory cytokines that recruit and activate neutrophils leading to inflammation. When these neutrophils undergo apoptosis and release their DNA the cycle is repeated and inflammation is maintained or increased. Thus, methods and reagents that block DNA binding to cytokine producing cells may therefore provide better treatment of CF patients than are currently available.

Although there have been several reports in the art that DNA could bind to cell surfaces (Bennett, 1993, *Antisense Res. Develop.* 3: 235-241; Bennett *et al.*, 1986, *J. Rheumatol.* 13: 679-685; Gabor & Bennett, 1984, *Biochem Biophys. Res. Commun.* 122:1034-1039; Hefeneider *et al.*, 1990, *J. Invest. Dermatol.* 94: 79S-84S; Bennett *et al.*, 1987, *J. Exp. Med.* 166: 850-863; Bennett *et al.*, 1991, *Clin. Exp. Immunol.* 86: 374-379; Bennett *et al.*, 1992, *Clin. Exp. Immunol* 90: 428-433; Bennett *et al.*, 1985, *J. Clin. Invest.* 76: 2182-2190; Hefeneider *et al.*, 1992, *Lupus* 1: 167-173; Hefeneider *et al.*, 1992, *Clin. Immunol. Immunopath.* 63: 245-251; Reid & Chalson, 1979, *Intl. Rev. Cytol.* 60: 27-52; Lerner *et al.*, 1971, *Proc. Natl. Acad. Sci. USA* 68: 1212-1216; Pancer *et al.*, 1981, *J. Immunol.* 127: 98-104; Meinke & Goldstein, 1974, *J. Molec. Biol.* 86:

757-773; Sudar *et al.*, 1986, *Cell. Molec. Biol.* 32: 87-91; Gasparro *et al.*, 1990, *Photochem & Photobiol.* 52: 315-321; Emlen *et al.*, 1988, *Amer. J. Pathol.* 133: 54-60), the art lacks an understanding of how cells mediate extracellular DNA binding. Thus, an understanding of the mechanisms by which eukaryotic cells, particularly mammalian cells, take up extracellular DNA would be important in improving a variety of biological processes.

SUMMARY OF THE INVENTION

The present invention relates to the cloning, expression and functional characterization of a mammalian DNA-R gene. The invention comprises nucleic acids having a nucleotide sequence of a novel mammalian DNA-R gene. The nucleic acids provided by the invention comprise a complementary DNA (cDNA) copy of the corresponding mRNA transcribed *in vivo* from the DNA-R genes of the invention. In a preferred embodiment, the mammalian DNA-R is a human DNA-R.. Also provided are the deduced amino acid sequence of the cognate proteins of the cDNAs provided by the invention, methods of making said cognate proteins by expressing the cDNAs in cells transformed with recombinant expression constructs comprising said cDNAs, and said recombinant expression constructs and cells transformed thereby.

This invention in a first aspect provides nucleic acids, nucleic acid hybridization probes, recombinant eukaryotic expression constructs capable of expressing the DNA-Rs of the invention in cultures of transformed cells, and such cultures of transformed eukaryotic cells that synthesize the DNA-Rs of the invention. In another aspect, the invention provides homogeneous compositions of the DNA-R proteins of the invention, homogeneous compositions of fragments of said DNA-R, most preferably a fragment comprising amino acids 1-575 of the DNA-R, as well as fusion proteins between the DNA-R or fragments thereof and, *inter alia*, epitope markers, and membrane preparations from cells expressing the DNA-R proteins of the invention, and also antibodies against and epitopes of the DNA-R proteins or fragments thereof of the invention. The invention in another aspect provides methods for making said homogenous preparations and membrane preparations using cells transformed with the recombinant expression constructs of the invention and expressing said DNA-R proteins thereby. Methods for characterizing the receptor and biochemical properties of these receptor proteins and methods for using these proteins in the development of agents having pharmacological uses related to the DNA-R of the invention are also provided.

In a first aspect, the invention provides a nucleic acid having a nucleotide sequence encoding a mammalian DNA-R. In a first preferred embodiment, the nucleic acid encodes a

human DNA-R. In this embodiment of the invention, the nucleotide sequence comprises 4351 nucleotides of human DNA-R cDNA comprising 3576 nucleotides of coding sequence, 601 nucleotides of 5' untranslated sequence and 177 nucleotides of 3' untranslated sequence. In this embodiment of the invention, the nucleotide sequence of the DNA-R is the nucleotide sequence depicted in Figure 1 (SEQ ID No:1). The sequence shown in Figure 1 will be understood to represent one specific embodiment of a multiplicity of nucleotide sequences that encode the human DNA-R amino acid sequence of 1192 amino acids (SEQ ID No.: 2) of the invention and that these different nucleotide sequences are functionally equivalent and are intended to be encompassed by the claimed invention. In addition, it will be understood that different organisms and cells derived therefrom express preferentially certain transfer RNAs (tRNAs) corresponding to subsets of the degenerate collection of tRNAs capable of encoding certain of the naturally-occurring amino acids, and that embodiments of the multiplicity of nucleotide sequences encoding the amino acid sequence of the human DNA-R protein of the invention that are optimized for expression in specific prokaryotic and eukaryotic cells are also encompassed by the claimed invention. Isolated nucleic acid derived from human genomic DNA and isolated by conventional methods using the human cDNA provided by the invention is also within the scope of the claimed invention. Finally, it will be understood that allelic variations of the human DNA-R; including naturally occurring and *in vitro* modifications thereof are within the scope of this invention. Each such variant will be understood to have essentially the same amino acid sequence as the sequence of the human DNA-R disclosed herein.

Mammalian DNA-R proteins corresponding to the human cDNA of the invention are a second aspect of the claimed invention. In a first embodiment, the mammalian DNA-R protein is a human DNA-R having a deduced amino acid sequence shown in Figure 1 (SEQ ID No.:2). In a second embodiment is provided said human DNA-R protein comprising a membrane preparation from a cell, most preferably a recombinant cell, expressing a nucleic acid encoding a human DNA-R of the invention.

As provided in this aspect of the invention is a homogeneous composition of a mammalian DNA-R having a molecular weight of about 150kD or derivative thereof that is a human DNA-R having an amino acid sequence shown in Figure 1 and identified by SEQ ID No.:2, said size being understood to be the predicted size of the protein before any post-translational modifications thereof. Also provided is a homogeneous composition of An amino-terminal fragment of the human DNA-R comprising amino acid residues 1-575 of the sequence identified as SEQ ID No.: 2. Species of the protein genetically engineered to lack the

transmembrane region of the DNA-R as described herein, and thereby providing soluble forms of the DNA-R of the invention, are also within the scope of this aspect of the invention and are provided herein.

5 This invention provides both nucleotide and amino acid probes derived from the sequences herein provided. The invention includes probes isolated from either cDNA or genomic DNA, as well as probes made synthetically with the sequence information derived therefrom. The invention specifically includes but is not limited to oligonucleotide, nick-translated, random primed, or *in vitro* amplified probes made using cDNA or genomic clone of the invention encoding a mammalian DNA-R or fragment thereof, and oligonucleotide and other
10 synthetic probes synthesized chemically using the nucleotide sequence information of cDNA or genomic clone embodiments of the invention.

It is a further object of this invention to provide such nucleic acid hybridization probes to determine the pattern, amount and extent of expression of the DNA-R gene in various tissues of mammals, including humans. It is also an object of the present invention to provide nucleic acid hybridization probes derived from the sequences of mammalian DNA-R genes of the
15 invention to be used for the detection and diagnosis of genetic diseases. It is an object of this invention to provide nucleic acid hybridization probes derived from the nucleic acid sequences of the mammalian DNA-R genes herein disclosed to be used for the detection of novel related receptor genes.

20 The present invention also includes synthetic peptides made using the nucleotide sequence information comprising the cDNA embodiments of the invention. The invention includes either naturally occurring or synthetic peptides which may be used as antigens for the production of DNA-R-specific antibodies, or useful as competitors of DNA-R molecules for nucleic acid binding, or to be used for the production of inhibitors of nucleic acid binding to
25 such DNA-R molecules.

The present invention also provides antibodies against and epitopes of the mammalian DNA-R molecules of the invention. It is an object of the present invention to provide antibodies that are immunologically reactive to the DNA-Rs of the invention. It is a particular object to provide monoclonal antibodies against these DNA-Rs. Hybridoma cell lines producing such
30 antibodies are also objects of the invention. It is envisioned that such hybridoma cell lines may be produced as the result of fusion between a non-immunoglobulin producing mouse myeloma cell line and spleen cells derived from a mouse immunized with a cell line which expresses antigens or epitopes of a mammalian DNA-R of the invention. The present invention also

provides hybridoma cell lines that produce such antibodies, and can be injected into a living mouse to provide an ascites fluid from the mouse that is comprised of such antibodies. It is a further object of the invention to provide immunologically-active epitopes of the mammalian DNA-R proteins of the invention. Chimeric antibodies immunologically reactive against the DNA-R proteins of the invention are also within the scope of this invention.

The present invention provides recombinant expression constructs comprising a nucleic acid encoding a mammalian DNA-R of the invention wherein the construct is capable of expressing the encoded DNA-R in cultures of cells transformed with the construct. A preferred embodiment of such constructs comprises a human DNA-R cDNA depicted in Figure 1 (SEQ ID No.:1), such constructs being capable of expressing the human DNA-R encoded therein in cells transformed with the construct. Also provided are recombinant expression constructs encoding fragments of said DNA-R, most preferably an amino-terminal fragment comprising amino acid residues 1-575 and fragments genetically engineered to lack the transmembrane domain of said DNA-R, thereby providing for production of soluble forms of the DNA-R. In alternative embodiments, the recombinant expression construct encodes a DNA-R fused to epitope sequences recognized by conventional antibodies known in the art. In each instance, the recombinant expression constructs of the invention are capable of expressing the human DNA-R encoded therein or fragment thereof in cells transformed with the construct.

The invention also provides prokaryotic and more preferably eukaryotic cells transformed with the recombinant expression constructs of the invention, each such cells being capable of and indeed expressing the mammalian DNA-R or fragment or epitope-modified species encoded in the transforming construct, as well as methods for preparing mammalian DNA-R proteins using said transformed cells.

The present invention also includes within its scope protein preparations of prokaryotic and eukaryotic cell membranes containing the DNA-R protein of the invention, or fragment or epitope-modified species thereof, derived from cultures of prokaryotic or eukaryotic cells, respectively, transformed with the recombinant expression constructs of the invention.

The invention also provides methods for screening compounds for their ability to inhibit, facilitate or modulate the biochemical activity of the mammalian DNA-R molecules of the invention, in particular nucleic acid binding thereto. In preferred embodiments, the methods of the invention relate to binding of DNA, particularly double-stranded DNA, and oligonucleotides. The methods of the invention are particularly directed towards identifying compounds that influence DNA or oligonucleotide uptake into cells expressing the DNA-R. In

preferred embodiments, the compounds identified by the methods of the invention influence DNA or oligonucleotide uptake by pinocytosis or endocytosis. In preferred embodiments, the compounds influence DNA or oligonucleotide uptake by increasing the amount of DNA or oligonucleotide that reaches the nucleus of the cell in a form that can be expressed therein.

5 Preferred compounds of the invention are identified by detecting increased uptake or increased expression of a gene, most preferably a reporter gene, encoded by said DNA. In preferred embodiments, cells transformed with a recombinant expression construct of the invention are contacted with such a compound, and the amount of DNA or oligonucleotide taken up by the cell, or the frequency or amount of gene expression, most preferably reporter gene expression,

10 in the cell is assayed.

Specific preferred embodiments of the present invention will become evident from the following more detailed description of certain preferred embodiments and the claims.

15 DESCRIPTION OF THE DRAWINGS

An understanding of the invention is facilitated by reference to the drawings.

Figure 1 is a schematic diagram illustrating cloning of the cDNA for the DNA-R of the invention. Antisera from a patient with systemic lupus erythroblastosis (SLE) and that inhibits cell surface DNA binding was used to screen λ gt11 library from peripheral blood mononuclear

20 cells. A positive clone (clone #88) containing an open reading frame (ORF) was obtained; the open reading frame remained open at the 5' end of the clone. Analysis of the nucleotide sequence of the clone identified a transmembrane region on the 3' end of the clone. A 731bp probe from was used to screen a λ gt11 cDNA library made from Raji cell line (human lymphoma cell line). Clone 97D42 which contained 462bp of additional 5' ORF sequence was

25 obtained from this clone. A modification of the polymerase chain reaction (5' random amplification of cDNA ends (RACE- PCR) was used to obtain the remainder of the 5' sequences from HeLa (human cervix carcinoma) and MOLT-4 (human lymphoblastic leukemia) cell lines. This sequence was compiled to produce an ORF of 3543bp that encoded a protein with a calculated molecular weight of 130.5kDa.

30

Figures 2A and 2B show Northern analysis of human cancer cell lines (Figure 2A) and human tissues (Figure 2B). A 442 bp DNA fragment (probe 11) from the 3' end of the gene coding for DNA-R was used as the radiolabeled probe for each blot.

5 Figure 3 is a schematic diagram of the human DNA-R of the receptor showing the location of the RING finger, zinc finger, proline rich and hydrophobic regions. An * denotes the N-linked glycosylation sites at amino acid positions 122, 394, 430, 451, 466, 468 and 1150.

10 Figure 4A is an alignment of conserved cysteines, histidines and aspartic acids of the C3HC3D RING finger in *Homo sapiens* ARD1 GTP-binding protein (Gene Bank Accession 422756), *H. sapiens* CART1 protein (Gene Bank Accession No. _____), *H. sapiens* SBBI03 hypothetical protein (Gene Bank Accession 5032071), *Caenorhabditis elegans* cDNA EST (Gene Bank Accession 3879246), *C. elegans* hypothetical 25.8 KD protein (Gene Bank Accession 2496825) *C. elegans* cDNA EST (Gene Bank Accession 3878739).

15 Figure 4B is an alignment of conserved cysteines and histidines of the C3H type zinc finger in *C. elegans* PIE-1 (Gene Bank U62896), *Drosophila melanogaster* DTIS 11 (Gene Bank U13397), *H. sapiens* TIS11B Buryrate response factors (EFT-Response factor) (Gene Bank X79066), *Saccharomyces cerevisiae* CTH1 Zinc finger protein (Gene Bank L42133).

20 Figure 5 is an analysis of DNA-R protein expression in mammalian cells. DNA-R was immunoprecipitated both as the native molecule and in an HA-tagged embodiment from stably-transfected human 293 cells (DNA-R/flu cells). Lane 1, lysate of 293 cells; lane 2, lysate of 293-DNA-R/flu cells; lanes 3-6, immunoprecipitation of 293-DNA-R/flu cell lysates with:
25 rabbit preimmune serum (lane 3) or anti-DNA-R (lane 4), control mouse monoclonal antibody (lane 5) or anti-HA (lane 6). Detected by Western blotting with rabbit anti-DNA-R IgH, immunoprecipitating rabbit IgG heavy chain.

30 Figure 6 shows the intracellular location of the DNA-R protein, associated with cell membranes in 293-MNAB/flu cells as detected by Western blotting with anti-HA (left half) or anti-DNA-R (right half). T, Triton X-100 whole-cell lysate; M, crude membrane fraction; C, cytosolic fraction.

Figure 7 illustrates immunofluorescence staining using anti-DNA-R and anti-transferrin antibodies on fixed, permeabilized A549 cells. A, Double stained with rabbit and sheep nonimmune sera. B, anti-DNA-R. C, anti-transferrin receptor. D, double stained for anti-DNA-R (red) and anti-transferrin receptor (green); colocalized staining appears yellow.

5

Figure 8A shows the results of antibody staining of 293 cell surface using polyclonal rabbit antisera raised against an amino-terminal fragment (aas 1-575) of the DNA-R of the invention. 293 cells were incubated with preimmune serum (black bars) or immune serum (white bars). Antibody binding was detected with FITC conjugated goat anti-rabbit IgG by flow cytometry. Each bar represents the geometric mean fluorescence intensity \pm sd (n=3, 10,000 viable cells in each analysis). The geometric mean fluorescence of the secondary antibody alone was $7.6 \pm .08$ (n=3).

10

Figure 8B shows antibody binding to the cell surface of Raji and 293 cells using antibodies raised against an amino-terminal fragment (1-575). Cells (Raji, gray bars; 293, white bars) were incubated with serial dilutions of rabbit antisera (#41 bleed 2) produced against the amino terminal portion (amino acids 1-575) of the DNA receptor. The cells were then incubated with FITC-goat anti-rabbit IgG and the fluorescence intensity measure by FACS. Fluorescence due to preimmune sera has been subtracted.

15

20

Figure 9A is a schematic diagram of the preparation of the soluble DNA-R protein from the full length DNA-R.

Figure 9B is an SDS-PAGE analysis of expression, affinity purification, and proteolysis of a fusion protein (termed GST/DNA-R) created between glutathione-S-transferase (GST) and an amino-terminal fragment (1-575) of the DNA-R of the invention. Lane 1, whole cell extract of *E. coli* expressing GST/DNA-R; lane 2, GST/DNA-R bound to glutathione (GSH)-sepharose; lane 3, site-specific proteolysis of GST/DNA-R while bound to GSH-sepharose; lane 4, eluate from GSH-sepharose following on-gel proteolysis of GST/DNA-R containing highly purified DNA-R (1-575) peptide.

25

30

Figures 10A through 10E show that GST/DNA-R but not control proteins immobilized on glutathione sepharose bind a fluorescently-labeled plasmid DNA (YOYO/pGEM4Z). Shown

in the figure is glutathione sepharose without immobilized protein (Figure 10A), or with immobilized GST (Figure 10B), GST/HST.1 (Figure 10C), GST/CBD (Figure 10D), or GST/DNAR (Figure 10E) incubated with YOYO/pGEM4Z for 30 min at 4° C. After washing, the fluorescence intensity was measured by FACS.

5

Figures 11A and 11B is the result of a "Southwestern" blot of nucleic acid binding to DNA-R. The experiments were performed by SDS-PAGE analysis, transfer to nitrocellulose and then incubated with (Figure 11A) or without (Figure 11B) biotinylated DNA, followed by binding to streptavidin conjugated with horse radish peroxidase (HRP) and reaction with colorimetric substrate. Absorbance was measured at 450 nm. These results showed that
10 GST/DNA-R but not GST bound biotinylated DNA.

Figure 12 shows the results of an enzyme linked immunosorbent assay (ELISA) in which purified DNA-R fragment (1-575) bound to immobilized DNA. DNA-R fragment (at concentrations of 0, 1, and 10 µ/mL) was incubated with immobilized plasmid DNA in ELISA plates. The plates were then incubated with anti-DNA antibodies at a 1:100 dilution, followed by secondary antibody conjugated to HRP and reaction with colorimetric substrate. Absorbance was measured at 450 nm.
15

Figure 13 shows that a DNA-R fragment comprising a zinc finger domain (at amino acids 416-435) participates in DNA binding by GST/DNA-R(1-575). Cysteine residues (at positions 416 and 431) were independently altered to serine (and termed C416S and C431S, respectively) or alanine (and termed C416A and C431A, respectively) and ELISA performed to evaluate DNA binding. The zinc finger cysteines at either 416 or 431 were altered to either a serine (C416S, C431S) or an alanine (C416A, C431A). Binding (100ng/well) of wild-type
20 (GST/DNA-R) or mutant GST-fusion proteins or GST alone to immobilized DNA was detected using anti-GST by ELISA. Data are the mean \pm s.d. of triplicate determinations.
25

Figure 14 illustrates results of experiments demonstrating that a soluble form of the DNA-R of the invention binds DNA with high affinity using a nitrocellulose filter binding assay. Soluble DNA-R (sDNA-R) at a concentration of 2 nM and labeled DNA (200 pM, 1×10^6 cpm/pmol) held constant and increasing concentrations of unlabeled DNA. Data are the mean \pm s.d. of triplicate determinations. Inset: Scatchard transformation of the binding data.
30

Figure 15 illustrates competitive binding of plasmid DNA with soluble DNA-R. All samples have 0.25nM (YOYO labeled) pGEM-DNA. The diagonal bars have varying amounts of soluble DNA-R added to block DNA binding. The horizontal line bar is control GST protein.

5

Figure 16 are the results of fluorescence activated cell sorting (FACS) experiments illustrating cell-surface binding of YOYO labeled plasmid DNA in A549 cells. Cells were incubated in the presence (dashed line) and absence (solid line) of 5 $\mu\text{g}/\text{mL}$ YOYO-labeled pGEM4Z DNA. The geometric mean fluorescence of untreated and treated cells is 13 and 34 respectively. The difference between the two values (21) is the increase in fluorescence intensity due to YOYO/pGEM4Z. This method of analysis is used for all YOYO/DNA binding analyses by FACS in subsequent figures.

10

Figures 17A and 17B are the results of FACS analysis of YOYO labeled plasmid DNA binding to A549 cells in the presence of excess unlabeled DNA. Cell surface binding of YOYO/DNA to A549 cells. In Figure 17A (left panel), A549 cells were incubated with YOYO/pGEM4Z in the presence (solid line) and absence (dashed line) of a 25-100 fold excess calf thymus DNA for 2 hr at 4°C. In Figure 17B (right panel), specific DNA binding to A549 cells is shown as the difference in fluorescence intensity of YOYO/pGEM4Z bound to the cells using the data from the left panel,. Data are the mean \pm SEM of 4-9 determinations.

15

20

Figure 18 shows the effects of trypsin treatment on DNA binding to A549 cells. A549 cells were incubated at 4° C for 30 min with YOYO/pGEM4Z (0.5 to 4 $\mu\text{g}/\text{ml}$). After incubation the cells were washed (white bars) or trypsinized and washed (gray bars), then the fluorescence intensity was measured by FACS. The data are the mean \pm SD of triplicate determinations. Trypsin treatment was found to remove most cell-surface DNA binding.

25

Figure 19 shows the results of experiments indicating that DNA binding to the DNA-R of the receptor is calcium dependent. B16 cells were incubated with 1 $\mu\text{g}/\text{mL}$ YOYO/pGEM4Z in PBS with (white bar) or without (gray bar) 1 mM CaCl_2 , then the fluorescence intensity was measured by FACS. The data are mean \pm SD of triplicate measurements.

30

Figure 20 shows a time course of plasmid DNA uptake by cells expressing the DNA-R of the invention. B16 cells were incubated at 37° C with YOYO/pEGFP-N1 (0.6 to 12 µg/mL) for 1 hr (solid circles), 3 hr (open circles), or 5 hr (solid squares). The cells were then trypsinized to remove DNA bound to the cell-surface, and fluorescence intensity was measured by FACS.

Figure 21 shows that excess unlabeled DNA blocks internalization of YOYO/DNA by A549 cells. Cells were incubated for 2 hr at 37° C with YOYO/pGEM4Z (1 to 25 µg/mL) in the presence (solid line) and absence (dashed line) of a 25- to 100- fold excess of calf thymus DNA. The cells were then washed and the fluorescence measured by FACS. Data are the mean ± SEM of 5 experiments.

Figure 22 illustrates that internalization of plasmid DNA in cells expressing the DNA-R of the invention is temperature dependent. B16 cells were incubated with YOYO/pGEM4Z (12 µg/mL) at 4° C (white bars) or 37° C (gray bars) for the indicated times. The cells were then trypsinized to remove cell-surface bound plasmid, washed, and the fluorescence measured by FACS. The data are the mean ± SD of triplicate determinations.

Figures 23A and 23B show that cell-surface DNA binding to the DNA-R of the invention is related to DNA uptake. 293 cells (solid circles) and G361 cells (open circles) were incubated for 3.5 hr with YOYO/pGEM4Z (at concentrations of from 0.3 to 2.5 µg/mL) at 4° C for binding (Figure 23A, left panel) or 37° C for uptake (Figure 23B, right panel). The cells were then washed (binding) or trypsinized and washed (uptake), and the fluorescence measured by FACS.

Figure 24 shows expression in B16 cells of green fluorescence protein (GFP) transgene from plasmid DNA. Cells were incubated with 12 µg/mL pEGFP-N1 for 6 hr. After 48 hr the cells were trypsinized and fluorescence measured by FACS. Control cells incubated without DNA and cells treated with pEGFP-N1 without carrier showed similar fluorescence, whereas cells incubated with pEGFP-N1 complexed with lipofectamine showed increased fluorescence.

Figure 25 show that nocodazole increases GFP transgene expression from plasmid DNA in A549 cells. A549 cells were incubated for 5 hr at 37° C with 25 µg/mL of pEGFP-N1 in the presence (dashed line) and absence (solid line) of 33 µM nocodazole. After 24 hr the cells were trypsinized and fluorescence measured by FACS.

5

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The term "mammalian DNA-R" as used herein refers to proteins consisting essentially of, and having substantially the same biological activity as, the protein encoded by the amino acid depicted in Figure 1 (SEQ ID No.:2). This definition is intended to encompass natural allelic variations in the disclosed DNA-R. Cloned nucleic acid provided by the present invention may encode DNA-R protein of any species of origin, including, for example, mouse, rat, rabbit, cat, and human, but preferably the nucleic acid provided by the invention encodes DNA-Rs of mammalian, most preferably human, origin.

The nucleic acids provided by the invention comprise DNA or RNA having a nucleotide sequence encoding a mammalian DNA-R. Specific embodiments of said nucleic acids are depicted in Figure 1 (SEQ ID No.:1), and include any nucleotide sequence encoding a mammalian DNA-R having an amino acid sequence as depicted in Figure 1 (SEQ ID No.: 2). Nucleic hybridization probes as provided by the invention comprise any portion of a nucleic acid of the invention effective in nucleic acid hybridization under stringency conditions sufficient for specific hybridization. Mixtures of such nucleic acid hybridization probes are also within the scope of this embodiment of the invention. Nucleic acid probes as provided herein are useful for isolating mammalian species analogues of the specific embodiments of the nucleic acids provided by the invention. Nucleic acid probes as provided herein are also useful for detecting mammalian DNA-R gene expression in cells and tissues using techniques well-known in the art, including but not limited to Northern blot hybridization, *in situ* hybridization and Southern hybridization to reverse transcriptase - polymerase chain reaction product DNAs. The probes provided by the present invention, including oligonucleotides probes derived therefrom, are also useful for Southern hybridization of mammalian, preferably human, genomic DNA for screening for restriction fragment length polymorphism (RFLP) associated with certain genetic disorders.

25
30

The production of proteins such as mammalian DNA-R from cloned genes by genetic

engineering means is well known in this art. The discussion which follows is accordingly intended as an overview of this field, and is not intended to reflect the full state of the art.

Nucleic acid encoding a DNA-R may be obtained, in view of the instant disclosure, by chemical synthesis, by screening reverse transcripts of mRNA from appropriate cells or cell line cultures, by screening genomic libraries from appropriate cells, or by combinations of these procedures, in accordance with known procedures as illustrated below. Screening of mRNA or genomic DNA may be carried out with oligonucleotide probes generated from the nucleic acid sequence information from mammalian DNA-R nucleic acid as disclosed herein. Probes may be labeled with a detectable group such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with know procedures and used in conventional hybridization assays, as described in greater detail in the Examples below. In the alternative, mammalian DNA-R nucleic acid sequences may be obtained by use of the polymerase chain reaction (PCR) procedure, using PCR oligonucleotide primers corresponding to nucleic acid sequence information derived from a DNA-R as provided herein. *See* U.S. Patent Nos. 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis.

Mammalian DNA-R protein may be synthesized in host cells transformed with a recombinant expression construct comprising a nucleic acid encoding the DNA-R nucleic acid, comprising genomic DNA or cDNA. Such recombinant expression constructs can also be comprised of a vector that is a replicable DNA construct. Vectors are used herein either to amplify DNA encoding a DNA-R and/or to express DNA encoding a DNA-R gene. For the purposes of this invention, a recombinant expression construct is a replicable DNA construct in which a nucleic acid encoding a DNA-R is operably linked to suitable control sequences capable of effecting the expression of the DNA-R in a suitable host.

The need for such control sequences will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator or enhancer sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. Amplification vectors do not require expression control domains. All that is needed is the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. *See*, Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Press: New York).

Vectors useful for practicing the present invention include plasmids, viruses (including phage and mammalian DNA and RNA viruses), retroviruses, and integratable DNA fragments

(i.e., fragments integratable into the host genome by homologous recombination). The vector can replicate the gene of interest and function independently of the host genome, or can, in some instances, integrate into the genome itself. Suitable vectors will contain replicon and control sequences which are derived from species compatible with the intended expression host.

5 Transformed host cells are cells which have been transformed or transfected with recombinant expression constructs made using recombinant DNA techniques and comprising nucleic acid encoding a DNA-R protein. Preferred host cells are HEK293 cells, COS-7 cells (Gluzman, 1981, *Cell* 23: 175-182) and Ltk⁻ cells. Transformed host cells may express the DNA-R protein, but host cells transformed for purposes of cloning or amplifying nucleic acid hybridization

10 probe DNA need not express the receptor. When expressed, the DNA-R of the invention will typically be located in the host cell membrane. Accordingly, the invention provides preparations of said cell membranes comprising the DNA-R protein of the invention, as well as purified, homogeneous preparations of the receptor protein itself. *See, Sambrook et al., ibid.*

Cultures of cells derived from multicellular organisms are a desirable host for recombinant DNA-R protein synthesis. In principal, any higher eukaryotic cell culture is useful,

15 whether from vertebrate or invertebrate culture. However, mammalian cells are preferred, as illustrated in the Examples. Propagation of such cells in cell culture has become a routine procedure. *See Tissue Culture*, Academic Press, Kruse & Patterson, editors (1973). Examples of useful host cell lines are human embryonic kidney (HEK) 293 cells, VERO and HeLa cells,

20 Chinese hamster ovary (CHO) cell lines, mouse Ltk⁻ cell lines and WI138, BHK, COS-7, CV, and MDCK cell lines. HEK293 cell, COS-7 cells and Ltk⁻ cells are preferred.

The invention provides homogeneous compositions of mammalian DNA-R protein produced by transformed eukaryotic cells as provided herein. Each such homogeneous composition is intended to be comprised of a DNA-R protein that comprises at least 75%, more

25 preferably at least 80%, and most preferably at least 90% of the protein in such a homogenous composition; in said homogeneous preparations, individual contaminating protein species are expected to comprise less than 5%, more preferably less than 2% and most preferably less than 1% of the preparation. The invention also provides membrane preparations from cells expressing mammalian DNA-R protein as the result of transformation with a recombinant

30 expression construct, as described herein. Also specifically provided by the invention are fragments of the DNA-R of the invention, most preferably DNA binding fragments thereof. In preferred embodiments, said fragments include soluble forms of the receptor lacking the transmembrane domain and an amino-terminal fragment (most preferably amino acids 1-575)

comprising zinc finger and RING sequence motifs known in the art to be related to DNA-protein binding.

Mammalian DNA-R proteins made from cloned genes in accordance with the present invention may be used for screening compounds that effect DNA binding to cells *in vivo* and *in vitro*, as more fully described herein, and that affect DNA uptake and expression of genes encoded thereby. For example, host cells may be transformed with a recombinant expression construct of the present invention, a mammalian DNA-R expressed in those host cells, and the cells or membranes thereof used to screen compounds for their effect on DNA binding. By selection of host cells that do not ordinarily express a DNA-R, pure preparations of membranes containing the receptor can be obtained.

The recombinant expression constructs of the present invention are useful in molecular biology to transform cells which do not ordinarily express a DNA-R to thereafter express this receptor. Such cells are useful as intermediates for making cell membrane preparations useful for receptor binding activity assays, which are in turn useful for drug screening. The recombinant expression constructs of the present invention thus provide a method for screening potentially useful drugs at advantageously lower cost than conventional animal screening protocols. While not completely eliminating the need for ultimate *in vivo* activity and toxicology assays, the constructs and cultures of the invention provide an important first screening step for the vast number of potentially useful drugs synthesized, discovered or extracted from natural sources each year. This utility thereby enables rational drug design of novel therapeutically-active drugs using currently-available techniques (*see* Walters, "Computer-Assisted Modeling of Drugs", *in* Klegerman & Groves, eds., 1993, Pharmaceutical Biotechnology, Interpharm Press:Buffalo Grove, IL, pp. 165-174).

The recombinant expression constructs of the present invention may also be useful in gene therapy. Cloned genes of the present invention, or fragments thereof, may also be used in gene therapy carried out homologous recombination or site-directed mutagenesis. *See generally* Thomas & Capecchi, 1987, *Cell* 51: 503-512; Bertling, 1987, *Bioscience Reports* 7: 107-112; Smithies *et al.*, 1985, *Nature* 317: 230-234.

Nucleic acid and oligonucleotide probes as provided by the present invention are useful as diagnostic tools for probing DNA-R gene expression in tissues of humans and other animals. For example, tissues are probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiographic or other detection techniques, to investigate native expression of this receptor or pathological conditions relating thereto. Further, chromosomes can be probed

to investigate the presence or absence of the corresponding DNA-R gene, and potential pathological conditions related thereto. Oligonucleotides, particularly antisense oligonucleotides, are also useful for decreasing expression of the DNA-R in cells that overexpress the receptor or whose expression is disadvantageous in a host organism, either generally or in specific tissues. An example of the latter instance is in airway epithelial cells and macrophages in lung tissues in cystic fibrosis patients, as set forth more fully herein.

The invention also provides antibodies that are immunologically reactive to the DNA-R protein or epitopes thereof provided by the invention. The antibodies provided by the invention may be raised, using methods well known in the art, in animals by inoculation with cells that express a DNA-R or epitopes thereof, cell membranes from such cells, whether crude membrane preparations or membranes purified using methods well known in the art, or purified preparations of proteins, including protein fragments and fusion proteins, particularly fusion proteins comprising epitopes of the DNA-R protein of the invention fused to heterologous proteins and expressed using genetic engineering means in bacterial, yeast or eukaryotic cells, said proteins being isolated from such cells to varying degrees of homogeneity using conventional biochemical methods. Synthetic peptides made using established synthetic methods *in vitro* and optionally conjugated with heterologous sequences of amino acids, are also encompassed in these methods to produce the antibodies of the invention. Animals that are useful for such inoculations include individuals from species comprising cows, sheep, pigs, mice, rats, rabbits, hamsters, goats and primates. Preferred animals for inoculation are rodents (including mice, rats, hamsters) and rabbits. The most preferred animal is the mouse.

Cells that can be used for such inoculations, or for any of the other means used in the invention, include any cell line which naturally expresses the DNA-R provided by the invention, or more preferably any cell or cell line that expresses the DNA-R of the invention, or any epitope thereof, as a result of molecular or genetic engineering, or that has been treated to increase the expression of an endogenous or heterologous DNA-R protein by physical, biochemical or genetic means. Preferred cells are mammalian cells, most preferably cells syngeneic with a rodent, most preferably a mouse host, that have been transformed with a recombinant expression construct of the invention encoding a DNA-R protein, and that express the receptor therefrom.

The present invention also provides monoclonal antibodies that are immunologically reactive with an epitope derived from a DNA-R of the invention, or fragment thereof, present on the surface of such cells or in membrane preparations thereof or used after varying degrees

of biochemical purification. Particularly useful are soluble fragments of the DNA-R of the invention, including for example species of the receptor genetically engineered to remove the transmembrane domain, and amino-terminal fragments, most preferably DNA binding fragments of the receptor. Such antibodies are made using methods and techniques well known to those of skill in the art. Monoclonal antibodies provided by the present invention are produced by hybridoma cell lines, which are also provided by the invention and are made by methods well known in the art.

Hybridoma cell lines are made by fusing individual cells of a myeloma cell line with spleen cells derived from animals immunized with cells expressing a DNA-R of the invention, as described above. The myeloma cell lines used in the invention include lines derived from myelomas of mice, rats, hamsters, primates and humans. Preferred myeloma cell lines are from mouse, and the most preferred mouse myeloma cell line is P3X63-Ag8.653. The animals from whom spleens are obtained after immunization are rats, mice and hamsters, preferably mice, most preferably Balb/c mice. Spleen cells and myeloma cells are fused using a number of methods well known in the art, including but not limited to incubation with inactivated Sendai virus and incubation in the presence of polyethylene glycol (PEG). The most preferred method for cell fusion is incubation in the presence of a solution of 45% (w/v) PEG-1450. Monoclonal antibodies produced by hybridoma cell lines can be harvested from cell culture supernatant fluids from *in vitro* cell growth; alternatively, hybridoma cells can be injected subcutaneously and/or into the peritoneal cavity of an animal, most preferably a mouse, and the monoclonal antibodies obtained from blood and/or ascites fluid.

Monoclonal antibodies provided by the present invention are also produced by recombinant genetic methods well known to those of skill in the art, and the present invention encompasses antibodies made by such methods that are immunologically reactive with an epitope of an amino acid receptor of the invention. The present invention also encompasses antigen-binding fragments, including but not limited to F_v , $F(ab)$ and $F(ab)_2$ fragments, of such antibodies. Fragments are produced by any number of methods, including but not limited to proteolytic or chemical cleavage, chemical synthesis or preparation of such fragments by means of genetic engineering technology. The present invention also encompasses single-chain antibodies that are immunologically reactive with an epitope of a DNA-R, made by methods known to those of skill in the art.

The present invention also encompasses an epitope of a DNA-R of the invention, comprised of sequences and/or a conformation of sequences present in the receptor molecule.

This epitope may be naturally occurring, or may be the result of chemical or proteolytic cleavage of a receptor molecule and isolation of an epitope-containing peptide or may be obtained by chemical or *in vitro* synthesis of an epitope-containing peptide using methods well known to those skilled in the art. The present invention also encompasses epitope peptides produced as a result of genetic engineering technology and synthesized by genetically engineered prokaryotic or eukaryotic cells.

The invention also includes chimeric antibodies, comprised of light chain and heavy chain peptides immunologically reactive to a DNA-R-derived epitope. The chimeric antibodies embodied in the present invention include those that are derived from naturally occurring antibodies as well as chimeric antibodies made by means of genetic engineering technology well known to those of skill in the art.

Nucleic acids encoding the receptor, the DNA-R and DNA-binding fragments thereof, are advantageously used to modulate expression or activity of the receptor in cells *in vivo* and *in vitro*. As provided herein, the DNA-R of the invention, particularly soluble embodiments thereof, can competitively bind DNA to reduce said binding to cells expressing the DNA-R. DNA binding to the DNA-R in certain cells, such as airway epithelial cells and macrophages in lung, is associated with the activation of inflammatory processes that cause a significant proportion of the morbidity and mortality associated with cystic fibrosis, chronic bronchitis and other chronic lung diseases. Thus, the invention provides a variety of methods for reducing said morbidity and mortality by interfering with DNA binding to cells in the lung. In one embodiment, soluble DNA-R species can be administered, most preferably by aerosol administration using formulations, excipients and vehicles well known in the art, directly to lung tissue, and competitive DNA binding achieved thereby. In alternative embodiments, antisense oligonucleotides can be delivered to lung tissue, most preferably by aerosol administration, and expression of the DNA-R in target cells of the lung repressed thereby. In further alternatives, antibodies, most preferably monoclonal antibodies, that specifically inhibit DNA binding to the DNA-R of the invention can be used to inhibit DNA binding to said lung cells.

The DNA-R of the invention, particularly soluble embodiments and DNA-binding fragments thereof, are also useful in treating other inflammation-associated diseases and conditions, including otitis media, septic arthritis and any bacterial or viral infection that causes inflammation by interaction with the DNA-R.

Additionally, the DNA-R of the invention can be used to screen compounds that modulate DNA binding, uptake and expression. Introduction of DNA, particularly DNA

encoding a desired gene, is a methodology well known in the art. However, DNA introduction methods have been developed empirically and without any understanding of the molecular bases of DNA uptake. Specifically, heretofore specific DNA binding to a DNA-R as disclosed herein and uptake thereby by endocytosis was unappreciated in the art. Identification of the DNA-R
5 of the invention thus provides a novel target for developing compounds and methods for increasing efficiency of DNA uptake and expression of genes encoded thereby.

Another advantageous method provided by this invention is the use of DNA-R expressed in tumor cells to facilitate delivery of DNA-binding anticancer drugs to tumor cells. Drugs such as Adriamycin (Doxorubicin) are in clinical use for the treatment of cancer patients. Enhanced
10 extracellular DNA uptake in tumor cells expression the DNA-R of the invention would facilitate uptake of such DNA-binding anticancer drugs by using extracellular DNA as a carrier of the drug into the cell. The association of the drug with the extracellular DNA might enable the drug to avoid active efflux produced in tumor cells, *inter alia*, by drug resistance mediators such as P-glycoprotein. Employing the same rationale as with gene transfer, the selective augmentation
15 of DNA binding receptors on tumor cells would enhance uptake of DNA-binding drugs and result in an increased therapeutic effect. In alternative embodiments, other diseases, such as malaria, can be treated in a similar fashion, based on the development of cell-surface DNA binding in red blood cells parasitized with the malarial parasite *Plasmodium falciparum*.

The Examples which follow are illustrative of specific embodiments of the invention,
20 and various uses thereof. They set forth for explanatory purposes only, and are not to be taken as limiting the invention.

EXAMPLE 1

Isolation of a Human Membrane-Associated DNA Receptor (DNA-R)

As described in the specification above, DNA binding to cells had been observed in the art, and the behavior of said binding suggested the existence of a DNA binding protein expressed at the cell surface. In order to isolate a novel DNA-binding protein from human cells, serum from a patient with systemic lupus erythematosus (SLE), treated to deplete the sera of
30 anti-DNA antibodies by multiple (6X) passages over a DNA sepharose column, was used to screen a λ gt11 cDNA expression library made from liposaccharide stimulated human monocytes. This serum has been shown to have anti-DNA receptor activities (defined by the

blocking of DNA binding to cells; Bennet *et al.*, 1992. *J. Clin. Invest.* 76: 2182-2190).

From approximately one million plaques screened with this sera, ten positive phage clones were identified and isolated according to the technique of Young and Davis (1983, *Proc. Natl. Acad. Sci. USA* 80: 1194-1198). The clones were grouped into two classes, based on Southern blot and Western blot analyses using eluted antibodies. Sequence analysis of the 1.4 kilobase (kb) insert of one clone (clone 88), which was highly reactive on Western blots with SLE serum, revealed an open reading frame that was open at the 5' end of the clone and contained a translation stop codon at the 3' end. This open reading frame coded for a 46.7 kiloDalton (kDa) protein fragment.

The full length cDNA for the putative DNA-R was obtained in segments from peripheral blood mononuclear cells, human Burkitt lymphoma cells (Raji; Accession No. CCL 86, American Type Culture Collection, Manassas, VA), human cervical carcinoma cells (HeLa; ATCC Accession No. CCL 2), and human lymphoblastic leukemia cells (MOLT-4; ATCC Accession No. CRL 1582). A 731 base pair (bp) DNA probe from clone 88 was used to screen a λ gt11 phage library from a Raji cell line (the library was obtained from Clonetech Labs, Palo Alto, CA). A 2409 bp clone which contained an additional 462 bp of 5' open reading frame (ORF) sequence was obtained from this screening. Additional sequence from the 5' extent of the cDNA was isolated using two variations of the 5' RACE (rapid amplification of cDNA ends) method. In the first 5' RACE method, single stranded DNA (ssDNA) was synthesized from HeLa cell mRNA using polyT primers and reverse transcriptase. A polyA tail was added to the 5' end of the ssDNA by terminal transferase. The single stranded cDNA was amplified using a gene specific primer and a polyT primer. This clone contained 753 bp of additional sequence 5' of the previously obtained sequence. The remainder of the 5' sequence of the cDNA was obtained from MOLT-4 cDNA by Marathon Race cDNA amplification (Clonetech Labs, Palo Alto, CA) according to the manufacturer's instructions. This procedure produced an additional 1290 bp clone consisting of 340 bp of ORF and a 950 bp 5' untranslated region. Combining the results from these screening and amplification experiments produced the predicted full length cDNA encoding the DNA-R of the invention.

A complete, full-length cDNA for the putative DNA receptor was cloned as a single RT-PCR product from MOLT-4 mRNA using oligonucleotide primers having the following sequence:

Primer 5': ACCCGAGCATGGATCCGCCACCATGGCTGTGCAGGCAGC (SEQ ID NO: 5)

and

Primer 3': GGTATCTAGATCCATGGTGTGGTCAC (SEQ ID NO: 6)

5 The complete sequence was 4351 nucleotides (SEQ ID NO: 1) in length with a defined open reading frame of 3576 nucleotides encoding a protein of 1192 amino acids (SEQ ID NO: 2). The isolation protocol is schematically illustrated in Figure 1.

10

EXAMPLE 2

DNA Receptor Gene Expression and Protein Sequence Analysis

Tissue-specific and cell line-specific expression patterns of its corresponding mRNA in various human tissues was analyzed by Northern blot analysis on RNA isolated from various tissues and cancer cell lines. The results of these experiments are shown in FIG. 2.

15 A panel of tissue samples was examined by Northern hybridization analysis performed under low stringency conditions, defined as hybridization at 42°C in 5X SSPE (0.75M NaCl, 0.05M NaH₂PO₄, 5mM EDTA), 10X Denhardt's solution (0.2% Ficoll, 0.2% polyvinylpyrrolidone, 0.2% bovine serum albumen), 100 µg/mL salmon sperm DNA, 2% SDS and 50% deionized formamide and 1-2 x 10⁶ cpm random-primed, ³²P labeled probe, followed
20 by washing in 0.1X SSC (15mM NaCl, 1.5mM trisodium citrate, 0.1% SDS). The blots were hybridized with a probe consisting of 442 bp of sequence from the 3' end of the coding sequence from the DNA-R gene to determine the distribution of receptor mRNA. This analysis revealed two major transcripts of 9.5 and 6.8 kb in all human tissues and cancer cell lines examined. Transcript expression was relatively abundant in spleen, testis, ovary, and small intestine. Several smaller transcript sizes were also observed in some of the tissues and cell lines
25 examined (FIG. 2).

A homology search against human genomic sequence placed the DNA receptor on chromosome 9q34 (GenBank Accession number AC007066, marker HIM9.89 on Contig CHR9.SL27). The genomic sequence, which covered 85% of the cDNA starting from the 5'
30 end, revealed the location of 16 complete exons and the beginning of a 17th exon. A BLAST search of the expressed tag sequence (EST) database indicated wide expression of this gene in normal human tissue (liver/spleen, prostate epithelial, germinal B cell, white adipose, pregnant uterus, fetal heart/liver and spleen) and in tumor and transformed human cells (Jurkat, HL60,

293, G361, B-cell lymphocytic leukemia, colon tumor, melanoma, and parathyroid tumor).

Figure 3 provides a schematic diagram of the structure of the DNA-R protein encoded by SEQ ID No. 1. Hydropathy analysis identified a 38 amino acid hydrophobic region near the carboxy terminus of the protein (amino acids 1133-1171) which is a potential transmembrane domain. Expression of a soluble species of this receptor by deleting these amino acids supported identification of this region as a transmembrane domain. In addition, seven consensus sites for α -linked glycosylation have been identified (amino acid positions 122, 394, 430, 451, 466, 468, and 1150) and there is a proline rich (20% of the residues are proline) region spanning amino acids 549-809 (FIG. 3). The calculated isoelectric point of the DNA receptor protein is 6.4. The BLAST search also identified two art-recognized amino acid sequence motifs in the DNA-R sequence: a C3HC3D Ring finger subtype located near the amino terminus (amino acids 14-50) and a C3H zinc finger located near the center of the protein sequence (amino acids 416-435). An alignment of several ring finger motifs is shown in FIG. 4A; DNA-R differs from the originally identified C3HC4 Ring finger motif by the replacement of the last cysteine with an aspartic acid. The alignment of the conserved cysteines and histidines of the C3H zinc finger motif is shown in FIG. 4B.

EXAMPLE 3

DNA Receptor Expression and Protein Expression Analyses

The DNA-R of the invention was produced recombinantly as follows. A *Bam*HI-*Hpa*I cDNA fragment containing the coding sequence for amino acids 1-1190 (*i.e.*, missing the two most carboxylterminal amino acids) of the DNA-R of the invention was cloned into pTriplFlu (obtained from J. Epstein, University of Pennsylvania, Philadelphia, PA). This vector contains a sequence encoding an epitope tag from the influenza hemagglutinin gene in triplicate inserted immediately 3' of the multiple cloning site of the parent vector, pcDNA3, and which are in-frame with the inserted DNA-R cDNA sequence. This vector was introduced into human 293 cells by transfection using Lipofectamine (Life Technologies, Gaithersburg, MD) according to the manufacturer's instructions. Transfected cells (293-DNA-R/flu) were selected by culturing in growth media (DMEM supplemented with 10% fetal calf serum, 2mM L-glutamine, 100U/mL penicillin and 100 μ g/mL streptomycin) supplemented with 500 μ g/mL G418.

In order to characterize DNA-R protein expression in mammalian cells, immunoprecipitation and Western blotting experiments were performed with protein extracts isolated from several cell lines using polyclonal antisera raised against an amino-terminal

fragment of the DNA-R of the invention, comprising amino acid residues 1-575.

Polyclonal antibodies were produced to a purified fragment of the DNA-R (comprising amino acids 1-575) using conventional techniques. Three female New Zealand White rabbits (Western Oregon Rabbit Company), weighing 2.3-3.0 kg, were injected subcutaneously with 50µg of the DNA-R peptide that was produced in bacteria as a GST fusion protein (described in Example 4) and purified from its fusion partner. The antigen was emulsified with Titre-Max (CytRx Corp., Norcross, GA) in a final volume of 0.5 mL. The rabbits were boosted 4 weeks later with 15 µg of antigen/Titre-Max mixture, again 2 weeks later, and were maintained on a once-a-month booster schedule thereafter. The rabbits were bled 7-10 days after each boost with antigen and the sera analyzed for reactivity to the immunizing antigen.

The polyclonal antisera obtained from the inoculated rabbits was used in Western blot analyses. A protein of Mr ~ 1.5 x 10⁵ was identified by the anti-DNA-R antibody in most cells tested (including 293, COS7, G361, HeLa, HRE605, MOLT-4, Raji, A549, B16). A protein with a similar mobility was detected in lysates of genetically-engineered human 293 cells (293-DNA-R/flu) that were stably transfected with an expression vector for a carboxy-terminal HA-tagged DNA-R (pDNA-R/flu). As shown in FIG. 5, this protein was detected by immunoprecipitation and/or Western blot analysis with either the rabbit polyclonal anti-DNA-R (1-575) antisera described above or with a mouse monoclonal antibody (anti-HA) specific for the carboxyl-terminal HA tag in the recombinantly-produced protein.

In order to determine cellular localization of the DNA-R protein, crude membrane fractions from recombinant 293-DNA-R/flu cells were examined by Western blot analysis with either anti-DNA-R or anti-HA antibodies. The results shown in FIG. 6 indicated that essentially all the DNA-R protein in those cells was associated with the membrane fraction. Indirect immunofluorescence on fixed, permeabilized cells showed anti-DNA-R staining was predominantly localized to the perinuclear region of the cell, although no nuclear staining was observed (FIG. 7). Double staining with anti-DNA-R and anti-transferrin receptor antibodies showed partial colocalization of the DNA-R and transferrin receptor, however the DNA-R did not colocalize with the transferrin receptor in peripheral endosomes (FIG. 7). These results indicate that extracellular DNA is taken up by cells expressing the DNA-R of the invention by endocytosis, and suggest that compounds that influence intracellular trafficking of molecules taken by endocytosis are useful for modulating the intracellular fate (such as degradation in lysosomes or transport to the cell nucleus) of extracellular DNA.

To determine if DNA-R is located on the cell surface, cells were incubated with anti-DNA-R (1-575) immune rabbit serum. Antibody binding was detected by flow cytometry with FITC labeled secondary antibodies to rabbit IgG. At all serum dilutions the fluorescence intensity of the cells incubated with immune serum was significantly higher than that of cells incubated with preimmune serum ($p < 0.003$) suggesting that DNA-R is expressed on the cell surface (FIG. 8).

These results demonstrated that the DNA-R protein, either natively expressed or expressed from the cloned cDNA of the invention, or genetically-engineered embodiments thereof, localized to cell membranes as predicted by the hydropathy plot of the carboxyl terminus.

EXAMPLE 4

Soluble DNA-R Fusion Protein Binds DNA with High Affinity

The capacity of the DNA-R of the invention to bind DNA, and particularly the capacity of a soluble form of the DNA-R protein to bind DNA (which would be useful for the development of a therapeutic agent as described more particularly below) was determined. For these experiments, a fusion protein between the amino terminal portion of the DNA receptor (amino acids 1-575), lacking the transmembrane region but containing both the RING and zinc finger domains, was produced using the pGEX vector system (Pharmacia, Kalamazoo, MI) for expression of glutathione-S-transferase (GST)-fusion proteins in *E. coli* and named GST/DNA-R (1-575). A schematic diagram of the production of this protein fragment and its structure relative to the full-length DNA-R of the invention is shown in Figure 9A. Polyacrylamide gel analysis of the production, proteolysis, and purification of the recombinant DNA-R peptide is shown in FIG. 9B. The calculated molecular weights of the GST/DNA-R fusion protein and the DNA-R peptide are 90 kDa and 63 kDa respectively.

The purified GST/DNA-R fusion protein was then examined for its ability to bind plasmid DNA. Three independent *in vitro* assays were used to assess DNA binding by the fusion protein. First, the ability of GST/DNA-R, bound to glutathione sepharose beads, to bind YOYO-labeled plasmid DNA was determined by incubation with 0.9 μg YOYO/DNA in 0.5 mL of medium. (YOYO-1 is an intercalating fluorochrome that is fluorescent only when bound to DNA, obtained from Molecular Probes, Eugene, OR.) Beads (3.5×10^5) and YOYO/DNA

were incubated for 30 minutes at 4°C, washed once and then fluorescence intensity analyzed by FACS. As seen in FIG. 10, the GST/DNA-R fusion protein was extremely efficient in binding DNA whereas purified GST protein alone and two additional, unrelated GST-fusion proteins (GST-CBD and GST-HST.1, gifts from Dr. Roland Kwok, Vollum Institute, Portland OR) failed to show any DNA binding capability. Following FACS analysis an aliquot of glutathione sepharose-bound protein from each sample used in the DNA binding assay was analyzed by SDS-PAGE followed by Coomassie blue staining. An approximately equivalent amount of each GST-fusion protein was shown to be present in each sample.

To further assess whether the GST/DNA-R fusion protein was a DNA-binding molecule, a Southwestern blot was performed. The purified GST/DNA-R fusion protein and GST protein alone were electrophoresed on a polyacrylamide gel, electrophoretically transferred to nitrocellulose and then probed with biotinylated DNA. DNA binding was visualized by addition of streptavidin conjugated with horse radish peroxidase (HRP) using conventional methods. As seen in FIG. 11, purified GST/DNA-R fusion protein (FIG. 11B, lane 1), but not GST protein alone (FIG. 11B, lane 2) bound biotinylated plasmid DNA. Other peptides seen to react with biotinylated DNA/streptavidin-HRP in the GST/DNA-R samples (FIG. 11B, lanes 1) probably represent partially degraded GST/DNA-R peptides and/or traces of contaminating bacterial proteins. Lanes in Figure 11A represent no added DNA.

Third, as a final assessment of the DNA binding ability of the purified DNA receptor fragment (amino acids 1-575) the ability of the purified peptide to bind to ELISA plates coated with plasmid DNA (VARELISA dsDNA kit, Pharmacia) was determined. Binding of the DNA receptor peptide was detected using the rabbit anti-DNA-R polyclonal antisera described Example 3. As shown in FIG. 12, purified DNA-R peptide bound to DNA coated plates when tested at both 1 µg/mL and 10 µg/mL. Negative controls not including the DNA-R fragment showed no reactivity.

These results demonstrate that the DNA receptor gene of the invention encodes a protein that specifically binds DNA, and that the DNA binding portion of the molecule resided in the protein fragment having amino acid sequence 1-575 of the native protein.

Having demonstrated that the protein encoded by the cloned cDNA of the invention bound DNA, the affinity of soluble GST-DNA-R for DNA was estimated using a nitrocellulose filter-binding assay. The assay was performed using cold DNA competition where known amounts of GST/DNA-R (2 nM) and labeled DNA (200 pM) were titrated with increasing

amounts of unlabeled calf thymus DNA. These results demonstrated that DNA binding to the DNA-R of the invention was saturable, consistent with its identification as a specific receptor. A Scatchard transformation of the data yield a $K_D \sim 4$ nM (FIG. 14).

To demonstrate that the binding of DNA by the soluble form of the DNA-R (amino acids 1-575) was not due to monospecific charge-related interactions, the role in DNA binding of the zinc finger domain at amino acids 416-435 was examined. Using site-directed mutagenesis, the codon for the conserved zinc finger cysteines at either amino acids 416 or 431 were altered to a codon for either alanine or serine. The mutagenized GST/DNA-R fusion proteins were expressed in *E. coli* and affinity purified on glutathione sepharose, then tested for their ability to bind to immobilized DNA by ELISA, all substantially as described above. Purified GST/DNA-R (1-575) fusion protein bound to ELISA plates coated with calf thymus DNA (Magiwell, United Biotech, Mountain View, CA), as shown in FIG. 13. Mutagenesis of either cysteine 416 or 431 reduced DNA binding to approximately 50% of the level observed for wild-type GST/DNA-R fusion protein, strongly suggesting that this zinc finger domain is involved in specific DNA binding FIG. 13.

These results demonstrated that DNA binding by the soluble DNA-R fragment is not simply a nonspecific charge related interaction, but rather is mediated by specific a DNA-binding motifs in the protein, including at least the zinc finger motif.

EXAMPLE 5

Soluble DNA-R Protein Inhibits DNA-Induced Cytokine Secretion and Blocks Binding of DNA to Cells

The presence of extracellular DNA in lung tissue of several chronic lung diseases, including cystic fibrosis, chronic bronchitis and bronchiectasis, causes or contributes to chronic inflammation of lung tissues with long-term pathological consequences. Extracellular DNA is known in the art to cause lung macrophages and other cells to release cytokines that mediate inflammation as part of the chronic symptomology of cystic fibrosis patients. As described in Example 2, the DNA-R protein of the invention is expressed in lung tissues, specifically in epithelial cells of the lung. This suggests that the DNA-R receptor protein of the invention is involved in inflammation by triggering the release of inflammation-causing cytokines. Thus, the ability of a soluble form of the DNA-R to bind DNA suggested that this protein fragment

could compete for binding extracellular DNA in cystic fibrosis patients and would be useful thereby as a therapeutic agent.

To determine if the soluble DNA-R fragment of the invention inhibits DNA-induced cytokine secretion, soluble DNA-R protein was examined for inhibition of CF-DNA-induced IL-6 release from J774 murine monocyte/macrophage cells in culture. In the absence of stimulating DNA, DNA-R did not induce IL-6 secretion (shown in Table I). DNA isolated and purified from a patient with cystic fibrosis (CF DNA) induced 611 pg/mL of IL-6 from J774 cells. When CF DNA was incubated first with DNA-R protein (10 ng/mL) and then added to J774 cells, the amount of IL-6 was reduced by 36% in the presence of the soluble DNA-R protein (10ng/mL). As a negative control, calf thymus DNA failed to induce detectable IL-6. To eliminate the possibility that cytokine release was caused by the presence of contaminating endotoxin, a Limulus amoebocyte assay was performed, and the CF DNA had < 0.25 ng/mL of contaminating endotoxin. In control experiments, this amount of LPS induced only 4 pg/mL of IL-6. In the second experiment (also shown in Table 1), contaminating endotoxin was removed from the soluble DNA-R, permitting the use of increased DNA-R concentrations. Soluble DNA-R protein (used in the range 10 ng/mL – 50 ng/mL) was incubated with J774 cells and 50 µg/mL of *E. coli* DNA. Cell-free supernatants were collected and IL-6 quantified by ELISA. In the absence of bacterial DNA soluble DNA-R did not induce IL-6 secretion. When bacterial DNA was added to the system, however, soluble DNA receptor protein inhibited IL-6 secretion in a dose-dependent manner (Table I).

TABLE I

Stimulus	Treatment	IL-6 (pg/mL)	% Inhibition
---	Medium	0	---
---	DNA-R (10ng/mL)	0	---
CF DNA (10µg/mL)	Medium	611	---
CF DNA (10µg/mL)	DNA-R (10ng/mL)	438	22
<i>E. coli</i> (10µg/mL)	Medium	1467	---
<i>E. coli</i> (10µg/mL)	DNA-R (10ng/mL)	945	36
<i>E. coli</i> (50µg/mL)	Medium	2390 ± 344	---
<i>E. coli</i> (50µg/mL)	DNA-R (10ng/mL)	1193 ± 128	50.1
<i>E. coli</i> (50µg/mL)	DNA-R (20ng/mL)	983 ± 212	58.9

Stimulus	Treatment	IL-6 (pg/mL)	% Inhibition
<i>E. coli</i> (50µg/mL)	DNA-R (50ng/mL)	652 ± 76	72.7
CT DNA ¹	Medium	0	---
LPS ²	Medium	4	---

¹ CT = calf thymus DNA

5 ² LPS = bacterial lipopolysaccharide (endotoxin)

To determine whether soluble DNA-R protein fragment was capable of preventing DNA binding to cells, J774 cells (5×10^5 cells) and YOYO labeled pGEM-DNA were incubated with either the soluble DNA-R protein fragment or control GST protein. Cells were incubated for 10 30 minutes at 4°C, centrifuged and washed twice with assay media, resuspended and incubated with 7-amino actinomycin D (7AAD) on ice for 20 minutes in order to assess viability. The samples were assessed for DNA binding by FACScan (Becton-Dickinson, Franklin Lanes, NJ). Results showed a dose-dependent inhibition of DNA binding to J774 cells (FIG. 15). Similar results were observed using human 293 cells. Additionally, the soluble DNA-R protein/DNA 15 complex does not bind to the cell surface. The soluble DNA-R protein bound to DNA and is effective at preventing the association of DNA with the cell surface.

These results indicate that the soluble DNA-R fragment provided by this invention is useful for inhibiting cytokine release, and inflammation consequent thereto, by competitively binding either bacterial or mammalian extracellular DNA and reducing the amount of such DNA 20 bound by cytokine-producing cells expressing the DNA-R of the invention.

EXAMPLE 6

DNA Binding to Cells Mediated by DNA-R

The experimental results disclosed above established that the soluble DNA-R fragment 25 comprising amino acids 1-575 of the DNA-R of the invention was capable of binding DNA. Further experiments were performed to characterize DNA binding to the receptor, particularly whether the native receptor protein was capable of binding extracellular DNA at the cell surface, and whether binding is consistent with a receptor-mediated process.

In these experiments, A549 human lung carcinoma cells were harvested from log-phase 30 cultures and treated with DNase and RNase to remove exogenous cell-surface bound nucleic acids. After treatment, the cells were washed with 10 mM EDTA and phosphate buffered saline

(PBS) to stop the action of DNase and RNase. The cells were then plated in V-bottom 96-well plates at 10^6 cells/well in PBS containing 1% fetal calf serum (FCS) and 1mM $\text{Ca}^{++}\text{Mg}^{++}$. YOYO-labeled pGEM4Z plasmid DNA was added at concentrations from 0 – 25 $\mu\text{g}/\text{mL}$ in 0.2 mL media containing 1% FCS and 1 mM $\text{Ca}^{++}\text{Mg}^{++}$. The cells plus labeled plasmid were incubated for 30 minutes at 4°C , to minimize internalization of plasmid DNA. Upon completion of the 30 minute incubation, the cells were washed with 2X in PBS containing 1% FCS and 1mM $\text{Ca}^{++}\text{Mg}^{++}$ and resuspended in 0.3 mL of PBS. Cells were then fixed in 1% formaldehyde and cell-surface binding of plasmid DNA quantified by FACS.

The results of these experiments are shown in FIG. 16. This representative FACS histogram demonstrates the A549 cell profiles seen when comparing cells incubated with either medium (FIG. 16, curve on the left) or cells incubated with 5 $\mu\text{g}/\text{mL}$ of YOYO/pGem4Z plasmid (FIG. 16, curve shifted to the right). The geometric mean of the intensity is used to describe the cell populations. In this example, the geometric mean of the A549 cell population, treated with medium only, was 13 and increased to 34 when incubated with YOYO-labeled plasmid DNA.

A binding curve for A549 cells was then generated using a range of plasmid DNA from 0-25 $\mu\text{g}/\text{mL}$ (FIG. 17A and 17B). The Y-axis of the graph in Figure 17A represents the geometric mean of the fluorescence intensity of the cell populations in the graph. Cell surface binding of plasmid DNA to A549 cells began to show saturation at approximately 10 $\mu\text{g}/\text{mL}$ of DNA. Treatment of cells with a 25-100 fold excess of unlabeled DNA significantly blocked the binding of YOYO/DNA to the cell surface (FIG. 17A). The specific cell-surface binding to A549 cells, represented as the difference between total binding seen with excess unlabeled DNA, shows a binding curve with a characteristic saturation profile (FIG. 17B).

Also examined were the cell-surface plasmid DNA binding profiles for a variety of tumor cell lines, including B16 murine melanoma cells, MOLT-4 human lymphoblastic leukemia cells, and the human Raji Burkitt lymphoma cells. In all cells examined, cell-surface DNA saturable binding profiles were obtained, consistent with a receptor-mediated mechanism of binding. Under optimal DNA binding conditions the percent of cells in the population capable of binding DNA above the background level as detected by FACS, ranged from greater than 70% (S49, DHL-6, MOLT-4) to less than 10% (D10.S, HUT-78, K562 and G361).

TABLE II

Cell Type	% Cells binding DNA	Lineage
S49	98	Murine T-cell lymphoma
MOLT-4	79	Human lymphoblastic leukemia
DHL-6	70	Human B-cell
A549	55	Human lung carcinoma
Dami	44	Human leukemia
B16	32	Murine leukemia
B9	21	Murine plasmacytoma
COS-7	20	African green monkey kidney cell
HBE014	20	Human bronchial epithelial cell
MO-7	16	Human leukemia
NOR-10	16	Murine muscle
J558	15	Murine plasmacytoma
RAJI	15	Human Burkitt lymphoma
HeLa	12	Human cervical cancer
SW480	12	Human colon adenocarcinoma
HUT-78	7	Human cutaneous T-cell lymphoma
K562	5	Human myelogenous leukemia
D10.S	3	Murine T cell
G361	3	Human malignant melanoma
Spleen	80	Normal mouse spleen cells

To determine if DNA binding is mediated by a cell-surface protein, the experiments were performed substantially as described after cells were treated with trypsin. Cell-surface DNA-binding of plasmid DNA on A549 cells was significantly inhibited by treatment of cells with trypsin after binding with YOYO-labeled DNA at 4°C (FIG. 18).

Finally, the effect of divalent cations on cell surface DNA binding was examined, using B16 melanoma cells. These studies demonstrated a four-fold increase in fluorescence intensity when Ca⁺⁺ is added to the binding media (FIG. 19).

These results indicate that the DNA-R protein of the invention mediates cell surface binding of extracellular DNA in mammalian cells.

5

EXAMPLE 7

Internalization of Extracellular DNA into Cells Expressing DNA-R

The experiments described in Example 6 established that extracellular DNA was specifically bound to the DNA-R of the invention. Internalization of DNA into cells by the receptor was characterized using the following assay.

10

YOYO-labeled plasmid DNA was used to examine the kinetics of plasmid DNA internalization. The plasmid used in these assays was pEGFP-N1, encoding green fluorescent protein (Clontech, Palo Alto, CA). The assay required that cell surface binding of labeled DNA be distinguished from internalized plasmid DNA. This was accomplished by treatment of cells with trypsin to remove cell-surface proteins after incubation with plasmid DNA. This procedure permitted cell surface-bound plasmid DNA to be distinguished from internalized plasmid DNA, since trypsin treatment abolished cell surface bound DNA but not internalized plasmid DNA. In this assay, cells were plated in 24 well plates and incubated in culture media for 24h. Media was then removed and various concentrations (0-25 $\mu\text{g/mL}$) of YOYO-labeled pEGFP-N1 plasmid DNA were added. The cells plus plasmid DNA were incubated for various times (0.5 to 5 hours) at 37°C. Thereafter, the media was removed, cells were treated with trypsin, washed, and then fixed with 1% formaldehyde. FACS analysis was used to quantify fluorescence intensity.

15

20

25

30

B16 murine melanoma cells were examined for internalization of YOYO/DNA using the above protocol after incubation for 1, 3, and 5 hours (FIG. 20). Internalization of pEGFP-N1 were found to be both dose- and time-dependent. An increasing amount of internalized plasmid DNA was seen with increasing dose of DNA and increasing time of incubation. Internalization of plasmid DNA by A549 cells was evaluated both with and without pre-treatment with unlabeled DNA. This assay was repeated with A549 cells, and similar results were obtained (FIG. 21). Moreover, pre-treatment of the A549 cells with a 25-100-fold excess of unlabeled calf thymus DNA significantly inhibited subsequent internalization of plasmid DNA (FIG. 21). Similar inhibition of internalization by pre-treatment of cells with excess unlabeled DNA was observed using a number of other cell lines (including B16, Raji, and

MOLT-4). This demonstration of saturable DNA binding and internalization indicates that the cell-surface DNA receptor of the invention mediates internalization of extracellular plasmid DNA.

Internalization of plasmid DNA was also observed to be a temperature-dependent process. Treatment of B16 cells at 4°C significantly inhibited the amount of plasmid DNA that was internalized as compared to cells maintained at 37°C (FIG. 22).

In order to ascertain whether the amount of DNA-R expressed on the cell surface influences the extent of extracellular DNA binding or DNA internalization, binding and internalization of plasmid DNA was compared in two cell lines: human melanoma G361 and the 293 human cell line. The G361 cells bound relatively low amounts of DNA, while 293 cells bound larger amounts of plasmid DNA as assessed by the cell-surface DNA binding assay (FIG. 23). Consistent with the binding results were the results obtained in these cells for DNA internalization, which showed that G361 cells internalized less plasmid DNA than 293 cells (FIG. 23). These data are consistent with identification of the DNA-R of the invention as a cell surface DNA receptor protein.

EXAMPLE 8

Gene Expression of DNA Internalized by DNA-R

Conditions for transgene expression using DNA internalized by the DNA-R of the invention were developed.

The experiments described above established plasmid DNA concentrations that saturated cell-surface binding. Used the pEGFP-N1 plasmid coding for green fluorescent protein (GFP), which was used because GFP remains exclusively intracellular. FACS analysis was used to quantify GFP expression. In this assay, cells (1.25×10^5 /well) were plated in 24-well plates and incubated overnight under mammalian cell culture conditions. On the next day, media was removed and the cells incubated for 3 hours at 37°C in 5% CO₂ with plasmid DNA in 0.3 mL of growth medium. DNA was then removed and fresh medium added to the cells. In some cases 0.3mL of growth medium was added to cells without removing the DNA. After 24-72 hours further incubation media was removed and cells washed once and then fixed with formaldehyde. Fluorescence intensity in the fixed cells was determined by FACS. However, no GFP expression was detected, even when using several different concentrations of pEGFP-N1 plasmid and incubation times. This result was consistently obtained, using a variety of cell

lines (A549, B16, Raji), incubation times (24-72 hours), and ranges of plasmid DNA concentrations (0.1 to 100 µg/mL). This result was obtained using cell lines that bind relatively higher levels of DNA on their cell-surface and those that bind lower levels of DNA. In positive controls, pEGFP-N1 plasmid was delivered by liposomes (Lipofectamine, Gibco-BRL, Gaithersburg, MD) and resulted in significant GFP fluorescence within 24 hours. Representative data using the B16 cell line incubated with either pEGFP-N1 alone or pEGFP-N1 delivered by liposomes shows the difference in GFP expression between these two techniques (FIG. 24).

In view of these results, the experiments were repeated with A549 cells in the presence of nocodazol, a microtubule inhibitor. Use of this inhibitor was indicated because one possible explanation of the unsuccessful experiments is that the DNA internalized by the DNA-R of the invention had been degraded, and nocodazol treatment was expected to reduce the extent of such degradation. Treatment of A549 cells with nocodazol prior to incubation with pEGFP-N1 resulted in a significant increase in expression of GFP as compared to cells that were not treated with nocodazol and incubated with pEGFP-N1 (FIG. 25). Cells which were not treated with nocodazol failed to demonstrate detectable expression of GFP (FIG. 25).

These results indicated that uptake of extracellular DNA mediated by the DNA-R of the invention required additional treatment to result in expression of genes encoded therein, and the above assay provides a prototype of assays for identifying such compounds. In these assays, an amount of GFP-encoding plasmid DNA known to reliably produce detectable GFP expression is contacted with a mammalian cell expression the DNA-R of the invention at levels known to mediate efficient uptake of extracellular DNA. GFP gene expression is then assayed in the presence and absence of a test compound to detect increased gene expression in the presence of the compound.

It should be understood that the foregoing disclosure emphasizes certain specific embodiments of the invention and that all modifications or alternatives equivalent thereto are within the spirit and scope of the invention as set forth in the appended claims.

WHAT WE CLAIM IS:

1. A nucleic acid comprising a nucleotide sequence encoding a mammalian cell surface DNA receptor (DNA-R).
2. A nucleic acid according to Claim 1 wherein the mammalian DNA-R is a human DNA-R and the nucleotide sequence of the nucleic acid encodes an amino acid sequence identified by SEQ ID No.:2.
3. A nucleic acid encoding a DNA-binding fragment of a mammalian DNA-R.
4. A nucleic acid according to Claim 3 wherein the mammalian DNA-R is a human DNA-R and the DNA-binding fragment comprises amino acids 1-575 of the amino acid sequence identified by SEQ ID No.: 2.
5. A nucleic acid encoding a soluble mammalian DNA-R.
6. A nucleic acid according to Claim 5 wherein the mammalian DNA-R is a human DNA-R having an amino acid sequence identified by SEQ ID No.: 2 wherein amino acids 1133-1171 are deleted therefrom.
7. A homogeneous composition of a mammalian DNA-R or derivative thereof having a molecular weight of about 150 kilodaltons and an amino acid sequence identified by SEQ ID No.:2.
8. A homogeneous composition of a DNA-binding fragment of a mammalian DNA-R or derivative thereof having a molecular weight of about 63 kilodaltons and an amino acid sequence identified by amino acids 1-575 of SEQ ID No.:2.
9. A homogeneous composition of a soluble mammalian DNA-R or derivative thereof having a molecular weight of about 145 kilodaltons and an amino acid sequence identified by SEQ ID No.:2 wherein amino acids 1133-1171 are deleted therefrom.
10. A recombinant expression construct comprising a nucleic acid having a nucleotide sequence encoding a mammalian DNA-R according to Claim 1, wherein the construct is capable of expressing the receptor in a transformed culture of eukaryotic or prokaryotic cells.
11. A recombinant expression construct according to Claim 10 wherein the mammalian DNA-R is a human DNA-R and the nucleotide sequence of the nucleic acid encodes an amino acid sequence identified by SEQ ID No.:2.

12. A recombinant expression construct comprising a nucleic acid having a nucleotide sequence encoding a DNA-binding fragment of a mammalian DNA-R according to Claim 3, wherein the construct is capable of expressing the receptor in a transformed culture of eukaryotic or prokaryotic cells.

13. A recombinant expression construct according to Claim 12 wherein the mammalian DNA-R is a human DNA-R and the DNA-binding fragment of the DNA-R is encoded by a nucleic acid encoding amino acids 1-575 of the amino acid sequence identified by SEQ ID No.:2.

14. A recombinant expression construct comprising a nucleic acid having a nucleotide sequence encoding a soluble mammalian DNA-R according to Claim 5, wherein the construct is capable of expressing the receptor in a transformed culture of eukaryotic or prokaryotic cells.

15. A recombinant expression construct according to Claim 14 wherein the mammalian DNA-R is a human DNA-R and the nucleotide sequence of the nucleic acid encodes an amino acid sequence identified by SEQ ID No.:2 wherein amino acids 1133-1171 are deleted therefrom.

16. A cell culture transformed with the recombinant expression construct of Claim 10, wherein the transformed cell culture expresses the mammalian DNA-R.

17. A cell culture transformed with the recombinant expression construct of Claim 11, wherein the transformed cell culture expresses the human DNA-R.

18. A cell culture transformed with the recombinant expression construct of Claim 12, wherein the transformed cell culture expresses the DNA-binding fragment of the mammalian DNA-R.

19. A cell culture transformed with the recombinant expression construct of Claim 13, wherein the transformed cell culture expresses the DNA-binding fragment of the human DNA-R.

20. A cell culture transformed with the recombinant expression construct of Claim 14, wherein the transformed cell culture expresses the soluble mammalian DNA-R.

21. A cell culture transformed with the recombinant expression construct of Claim 15, wherein the transformed cell culture expresses the soluble human DNA-R.

22. A method of screening a compound for modulating extracellular DNA binding to a cell that expresses a mammalian DNA-R, the method comprising the steps of:

- (a) selecting a host cell that expresses a mammalian DNA-R;

- (b) contacting the cell with detectably-labeled extracellular DNA in the presence and absence of the compound;
 - (c) assaying the cell for detectably-labeled DNA binding; and
 - (d) identifying compounds that modulate extracellular DNA binding by detecting a different amount of the detectable label bound to the cell in the presence of the compound than in the absence of the compound.
23. A method according to claim 22, wherein the host cell expresses an endogenous DNA-R.
24. A method according to claim 22, wherein the host cell expresses an exogenous DNA-R.
25. A method according to claim 22, wherein detectable label is a fluorescent label or a radiolabel.
26. A method according to claim 22, wherein the compound inhibits binding of extracellular DNA to the cell, wherein a decreased amount of extracellular DNA binds to the cell in the presence than in the absence of the compound.
27. A method according to claim 22, wherein the compound enhances binding of extracellular DNA to the cell, wherein an increased amount of extracellular DNA binds to the cell in the presence than in the absence of the compound.
28. A method according to claim 22, wherein the compound is an antibody or antigen-binding fragment thereof.
29. A method of screening a compound for modulating intracellular uptake of extracellular DNA, the method comprising the steps of:
- (a) selecting a host cell that expresses a mammalian DNA-R;
 - (b) contacting the cell with extracellular DNA that is detectably labeled in the presence and absence of the compound;
 - (c) assaying the cell for intracellular uptake of the detectable label; and
 - (d) identifying compounds that increase intracellular uptake of extracellular DNA by detecting a greater amount of the detectable label inside the cell in the presence of the compound when compared with the amount of detectable label inside the cell in the absence of the compound.
30. A method according to claim 29, wherein the host cell expresses an endogenous DNA-R.

31. A method according to claim 29, wherein the host cell expresses an exogenous DNA-R.
32. A method according to claim 29, wherein detectable label is a fluorescent label or a radiolabel.
33. A method according to claim 29, wherein the compound inhibits uptake of extracellular DNA to the cell, wherein a decreased amount of extracellular DNA is detected intracellularly in the presence than in the absence of the compound.
34. A method according to claim 29, wherein the compound enhances uptake of extracellular DNA to the cell, wherein an increased amount of extracellular DNA is detected intracellularly in the presence than in the absence of the compound.
35. A method according to claim 29, wherein the compound is an antibody or antigen-binding fragment thereof.
36. A method of screening a compound for modulating uptake or expression of extracellular DNA, the method comprising the steps of:
- (a) selecting a host cell that expresses a mammalian DNA-R;
 - (b) contacting the cell with extracellular DNA encoding a gene whose expression can be detected in the presence and absence of the compound;
 - (c) assaying the cell for expression of the gene encoded by the extracellular DNA; and
 - (d) identifying compounds that increase uptake or expression of extracellular DNA by detecting increased expression of the gene in the presence of the compound when compared with expression of the gene in the absence of the compound.
37. A method according to claim 36, wherein the host cell expresses an endogenous DNA-R.
38. A method according to claim 36, wherein the host cell expresses an exogenous DNA-R.
39. A method according to claim 36, wherein detectable label is a fluorescent label or a radiolabel.
40. A method according to claim 36, wherein the compound inhibits uptake of extracellular DNA to the cell, wherein a decreased amount of extracellular DNA is detected intracellularly in the presence than in the absence of the compound.

41. A method according to claim 36, wherein the compound enhances uptake of extracellular DNA to the cell, wherein an increased amount of extracellular DNA is detected intracellularly in the presence than in the absence of the compound.

42. A method according to claim 36, wherein the compound inhibits expression of a gene encoded by said extracellular DNA, wherein a decreased amount of the gene product encoded by the extracellular DNA is detected in the presence than in the absence of the compound.

43. A method according to claim 36, wherein the compound enhances uptake of extracellular DNA to the cell, wherein an increased amount of the gene product encoded by the extracellular DNA is detected in the presence than in the absence of the compound.

44. A method according to claim 36, wherein the compound is an antibody or antigen-binding fragment thereof.

45. A method according to claim 22, wherein the cell is transformed with a recombinant expression construct encoding a mammalian DNA-R, wherein the cells of the transformed cell culture express the receptor.

46. A method according to claim 29, wherein the cell is transformed with a recombinant expression construct encoding a mammalian DNA-R, wherein the cells of the transformed cell culture express the receptor.

47. A method according to claim 36, wherein the cell is transformed with a recombinant expression construct encoding a mammalian DNA-R, wherein the cells of the transformed cell culture express the receptor.

48. A cell membrane preparation comprising a mammalian DNA-R or derivative thereof having a molecular weight of about 150 kilodaltons.

49. A cell membrane preparation according to claim 48 wherein the mammalian DNA-R is human DNA-R having an amino acid sequence identified by SEQ ID No.:2.

50. A method for inhibiting extracellular DNA binding to a eukaryotic cell expressing a mammalian DNA-R, the method comprising the step of contacting the cell with a soluble mammalian DNA-R.

51. A method according to claim 50, wherein the eukaryotic cell expresses a human DNA-R having an amino acid sequence identified as SEQ ID No.: 2.

52. A method according to claim 50, wherein the soluble DNA-R is a human DNA-R having an amino acid sequence identified by SEQ ID No.:2 wherein amino acids 1133-1171 are deleted therefrom.

53. A method for inhibiting extracellular DNA binding to a eukaryotic cell expressing a mammalian DNA-R, the method comprising the step of contacting the cell with a DNA-binding fragment of a mammalian DNA-R.

54. A method according to claim 53, wherein the eukaryotic cell expresses a human DNA-R having an amino acid sequence identified as SEQ ID No.: 2.

55. A method according to claim 53, wherein the DNA-binding fragment of DNA-R is a fragment of human DNA-R comprising amino acids 1-575 of the amino acid sequence identified by SEQ ID No.:2.

56. An antibody or antigen-binding fragment thereof that is immunologically specific for an epitope comprising a mammalian DNA-R.

57. An antibody according to claim 56, wherein the mammalian DNA-R is human DNA-R having an amino acid sequence identified as SEQ ID No.: 2.

58. An antibody according to claim 56 comprising a polyclonal antisera.

59. An antibody according to claim 56, wherein the antibody is a monoclonal antibody

60. A hybridoma cell line that produces an antibody according to claim 59.

61. An antibody according to claim 56, wherein the epitope comprises a DNA binding domain identified by amino acids 1-575 of SEQ ID No.: 2.

62. A method for inhibiting inflammation in a tissue caused by extracellular DNA binding to cells comprising said tissue, the method comprising the step of contacting said tissue with a therapeutically-effective amount of a soluble mammalian DNA-R.

63. A method according to claim 62 wherein the tissue is lung tissue.

64. A method according to claim 63, wherein the cells of the lung tissue are epithelial cells or macrophages.

65. A method according to claim 62, wherein the soluble mammalian DNA-R is human DNA-R having an amino acid sequence identified as SEQ ID No.: 2 wherein amino acids 1133-1171 are deleted therefrom.

66. A method for inhibiting inflammation in a tissue caused by extracellular DNA binding to cells comprising said tissue, the method comprising the step of contacting said tissue with a therapeutically-effective amount of a DNA-binding fragment of a mammalian DNA-R.

67. A method according to claim 66 wherein the tissue is lung tissue.

68. A method according to claim 67, wherein the cells of the lung tissue are epithelial cells or macrophages.

69. A method according to claim 66, wherein the DNA-binding fragment of a mammalian DNA-R is human DNA-R comprising amino acids 1-575 of an amino acid sequence identified as SEQ ID No.: 2.

70. A method for inhibiting inflammation in a tissue caused by extracellular DNA binding to cells comprising said tissue, the method comprising the step of contacting said tissue with a therapeutically-effective amount of an antibody or antigen-binding fragment thereof immunologically specific for a mammalian DNA-R.

67. A method according to claim 66 wherein the tissue is lung tissue.

68. A method according to claim 67, wherein the cells of the lung tissue are epithelial cells or macrophages.

69. A method according to claim 66, wherein the antibody is a monoclonal antibody immunologically specific for a DNA-binding fragment of a mammalian DNA-R is human DNA-R comprising amino acids 1-575 of an amino acid sequence identified as SEQ ID No.: 2.

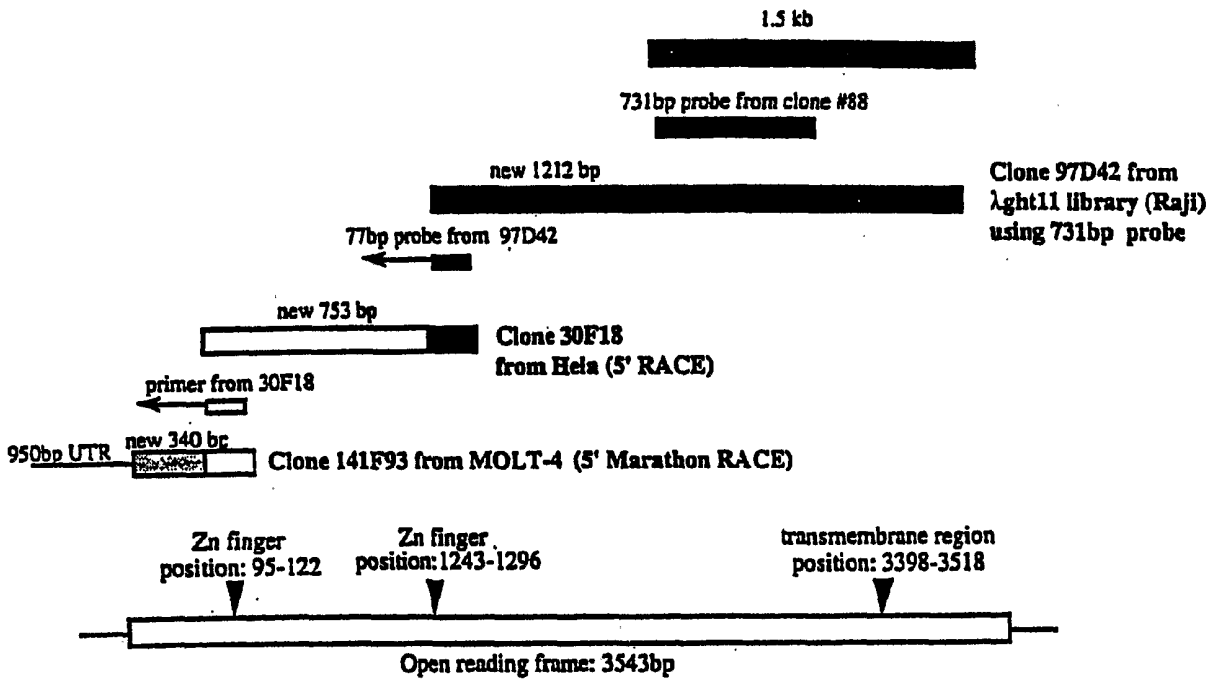


Figure 1

Figure 2.

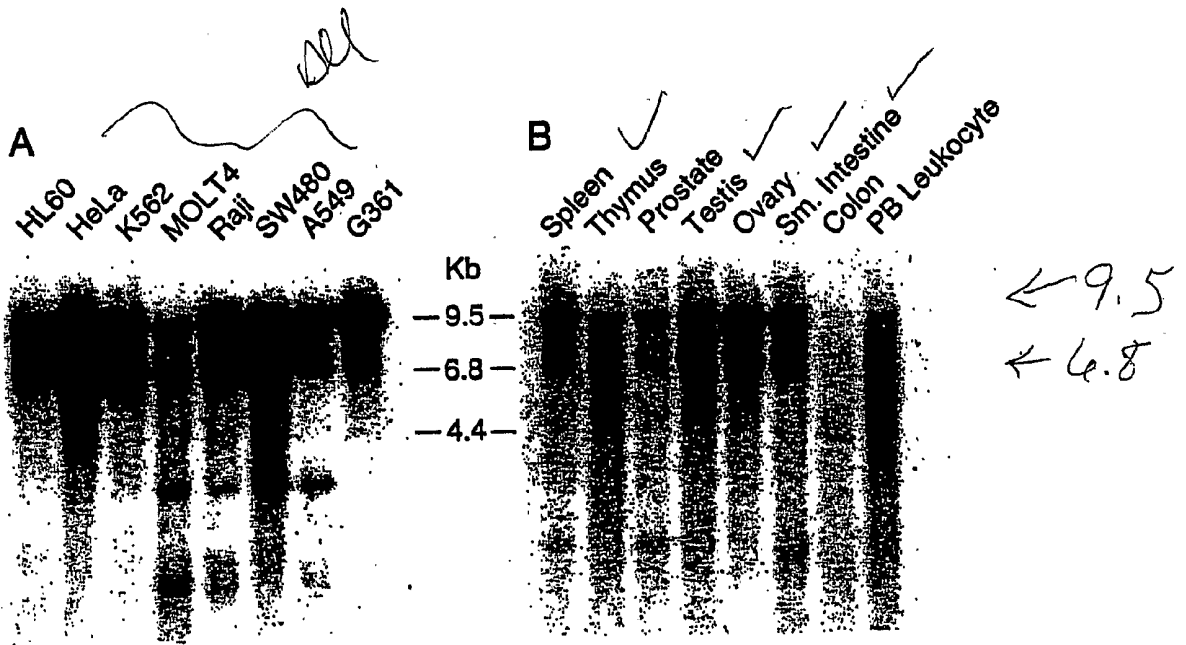


Figure 3

Open Reading Frame 1192 amino

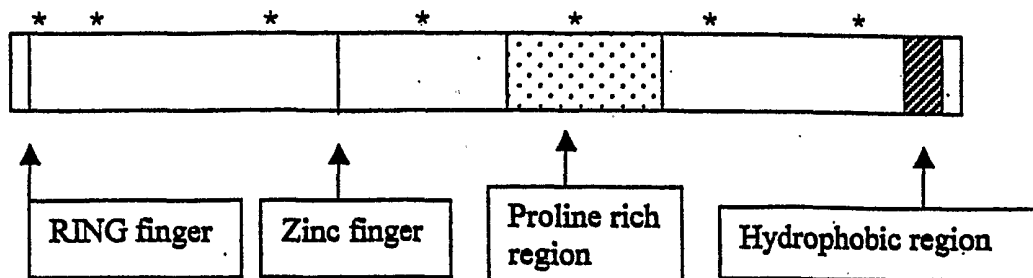


Figure 4

A. C3HC3D Ring finger homologies

MNAB	TEFLSCPIC---YNEFDENVHKPISLGCSETVCKTCLNKLHRKA-----CPFDQTAIN	58
H.sapiens ARD1	VKMLECGVCEDEVFSL--QGDKVPRLLLCGHTVCHDCLTRLPLHGR--AIRCPFDROVTD	80
H.sapiens CART1	KRRLCPLCGKEMREPVQVST-----CGHRFCDTCLQEFLESEG---VFKCFEDQLPLD	62
H.sapiens SBBT03	DEDLICPICSGVLEEFVQAFH-----CEHAFCNACITQWFSQQ---Q-TCEVDRSVVT	61
C.elegans cDNA EST 3879246	--YSECLVC---YQKFDENTRIPRVMDCGHTLCEFCINQIVKMAGCYSATCFPDRVRI-	182
C.elegans 25.8KD protein	-ENRSCFICTMEYSRTDKNLH-PIILNCGHNLCRSCINKL---TGNGIVKCFPDR----	199
C.elegans cDNA EST 3878739	QEVLCGSICNRHENE-----TFLPVSLICGHVICRKAEPENQTK----FCPHDQWKT	61
C3HC3D motif	-----C-C-----C-H-C-C-----C-D-----	

B. C3H Zinc finger homologies

MNAB	ETPQPQPNISKYKTSMCRDLRQGGCPRGTNCTFAHSQEELEKYRLRNK	448
C.elegans PIE-1	-----HTEYKTRRLCDAFRREGYCPYNDNCTYAHGQDELRVPRRRQE	136
D.melanogaster DTIS 11	---QPMNTSRYKTELCPFEAGECKYGEKQFAHGSHEL-----	166
H.sapiens TIS11B	-----SSRYKTELCPFEENGACKYGDKQFAHGIHEL-----	181
S. cerevisiae CTH1	QLPQLVNKTLTKTELCESTIKGYCKYGNKQFAHGLNELK-----	235
C3H motif	-----YKTELC-----C-----C-H-----	

Figure 5

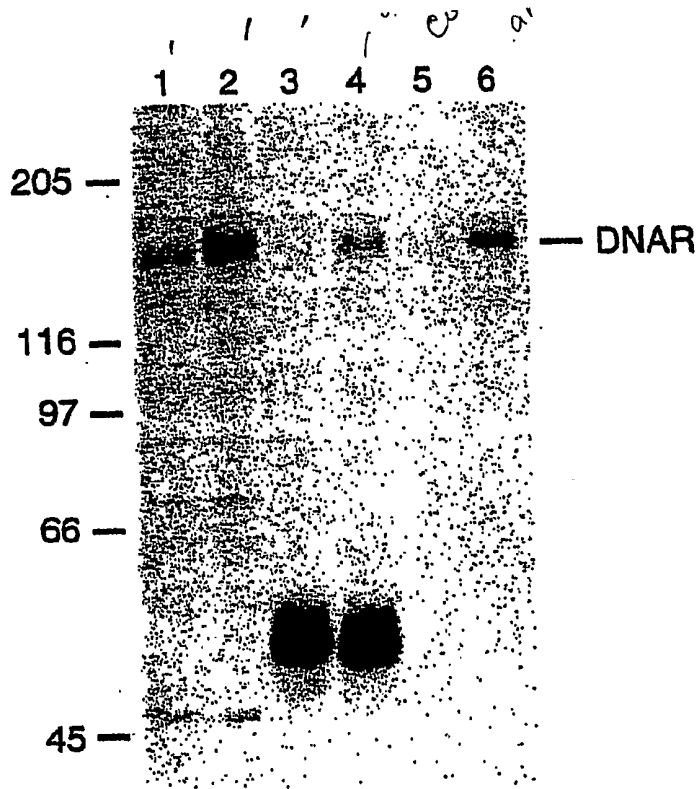
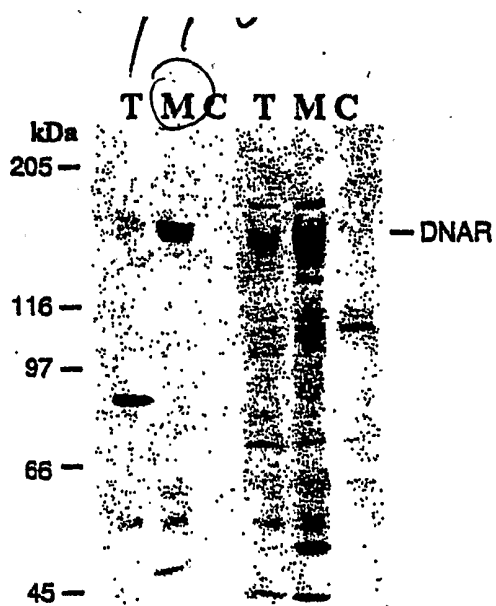


Figure 6



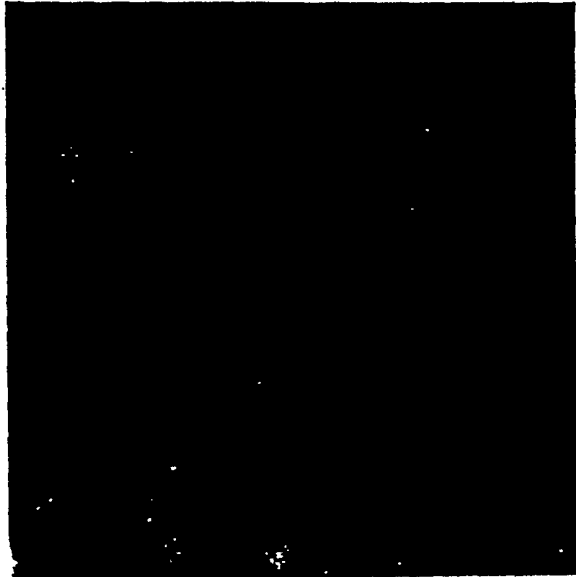


Figure 8

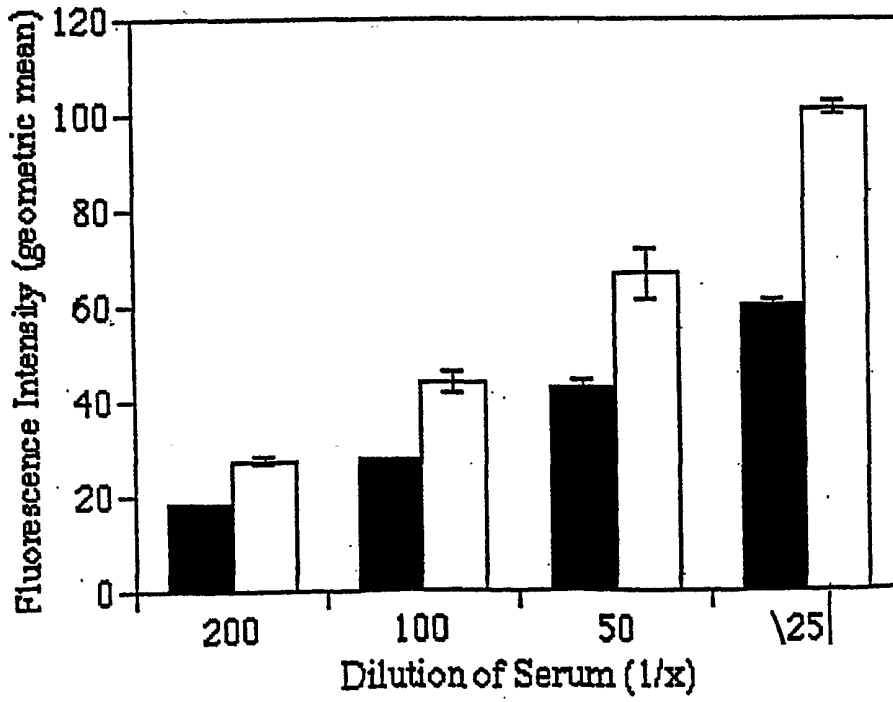


Figure 9 A

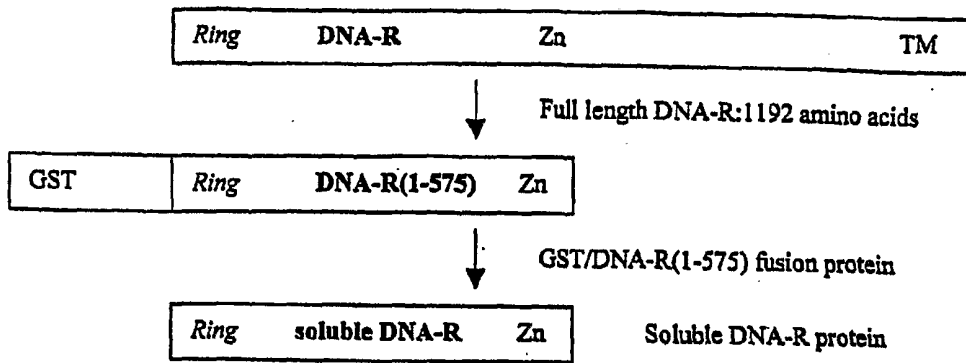
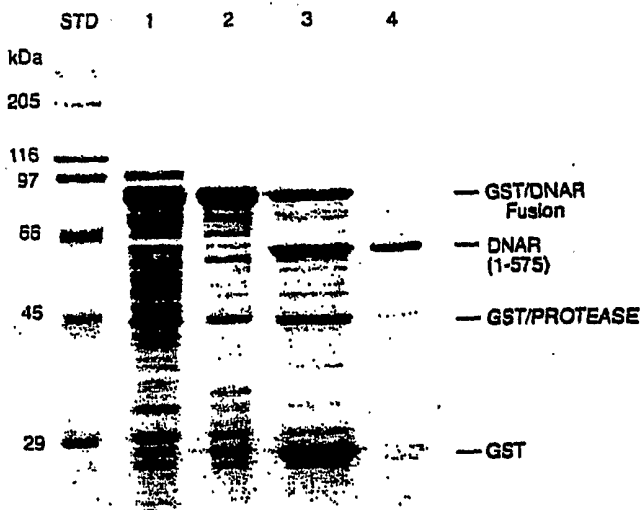


Figure 9 B



10/26

Figure 10

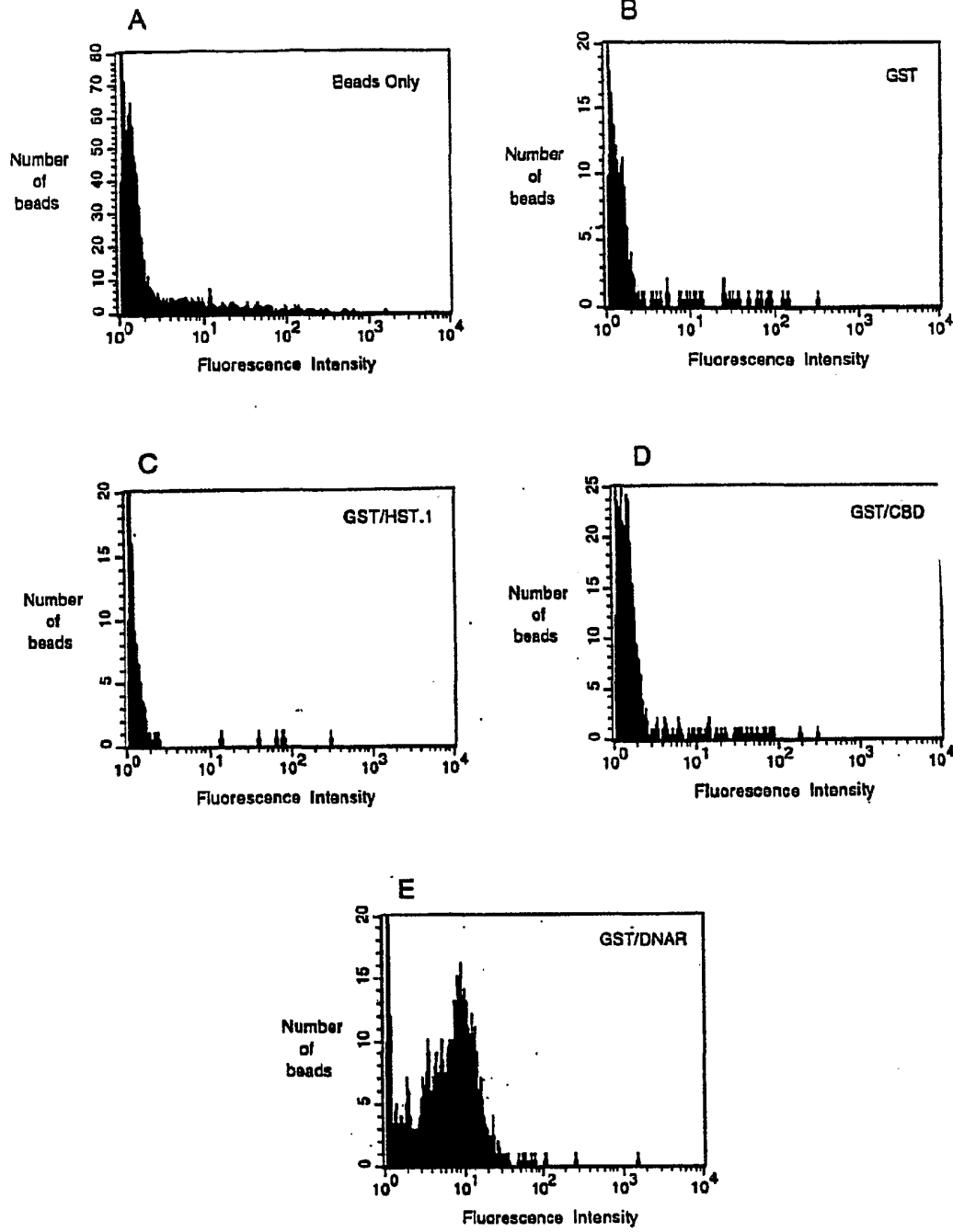


Figure 11

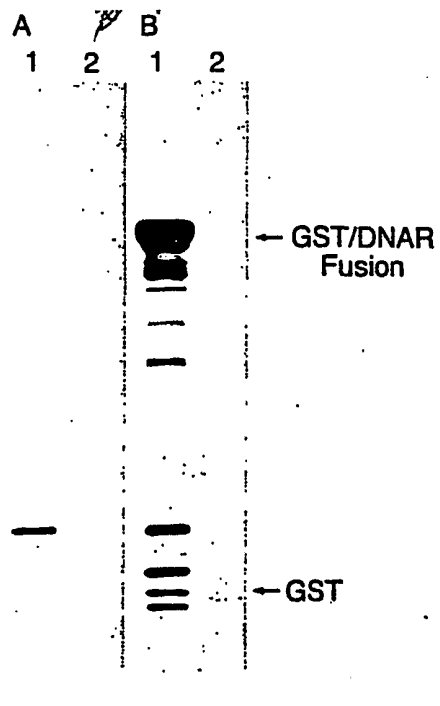


Figure 12

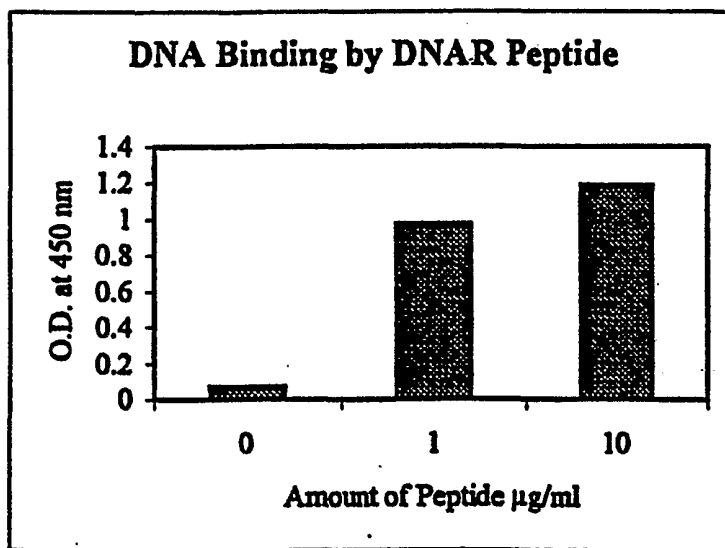


Figure 13.

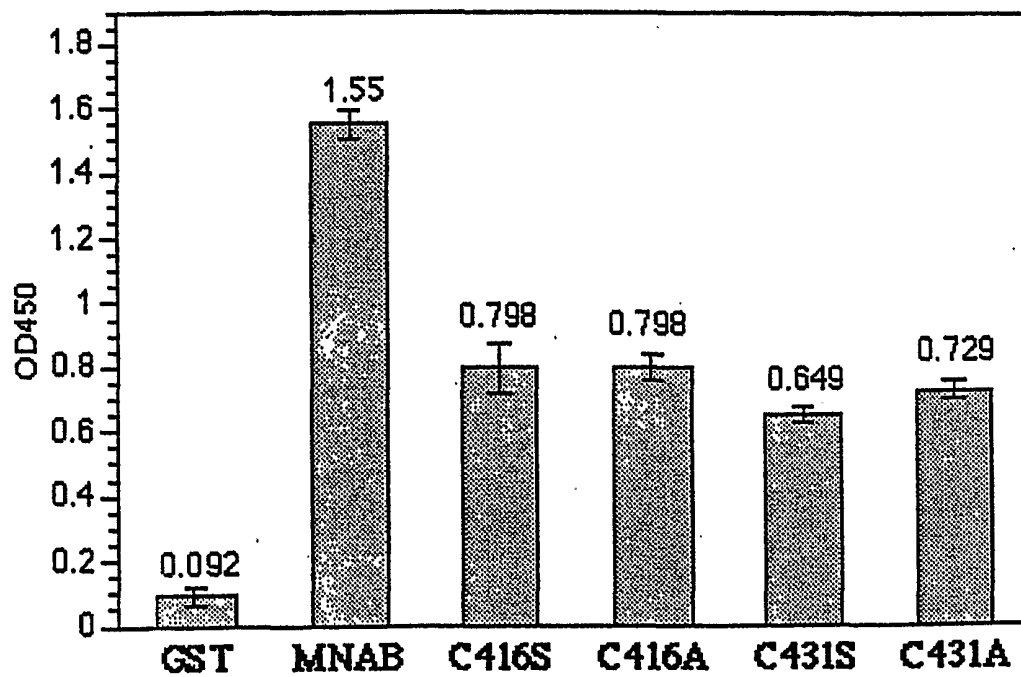


Figure 1A

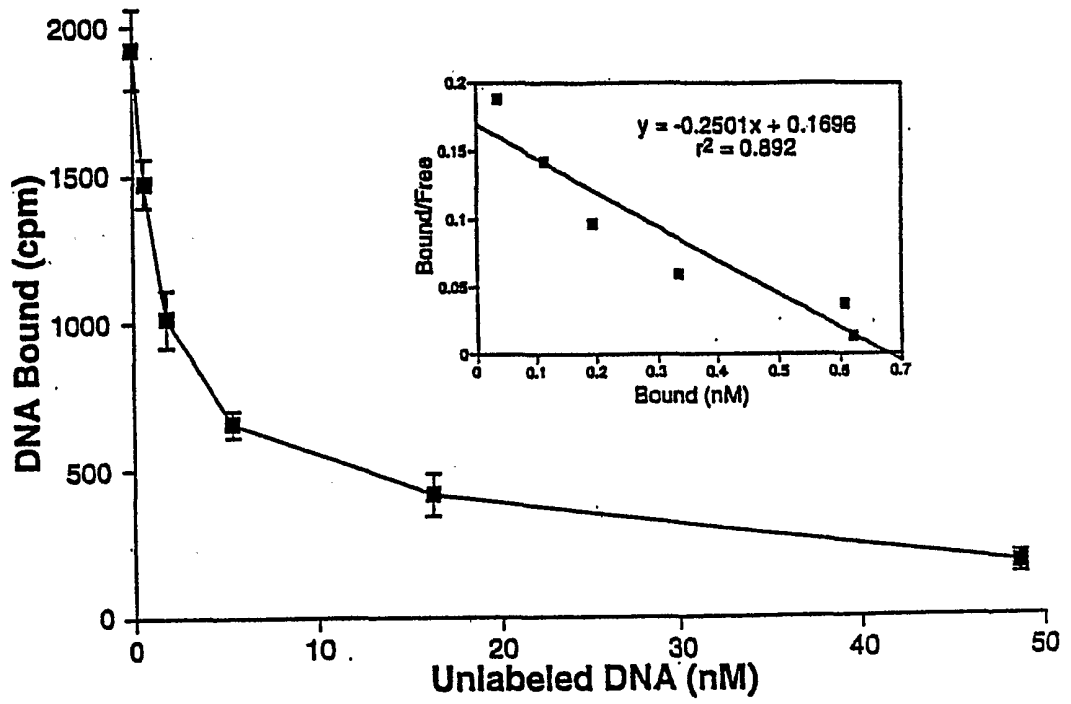


Figure 15

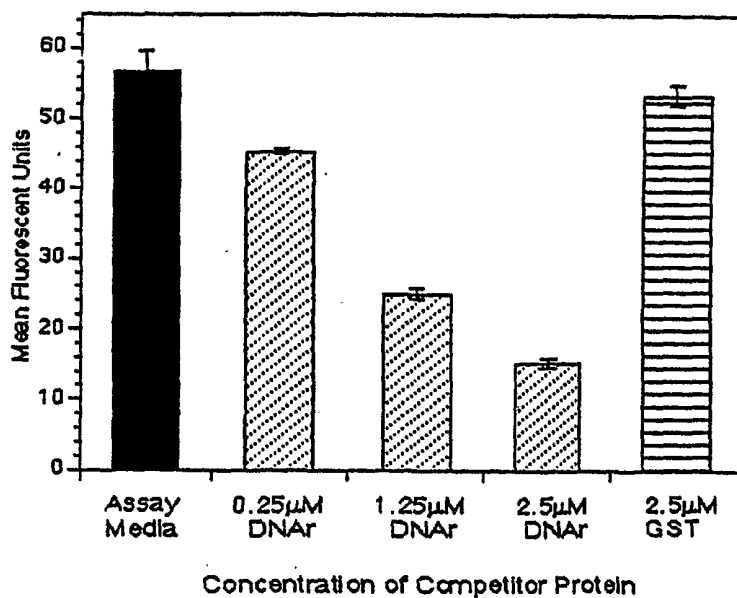


Figure 16

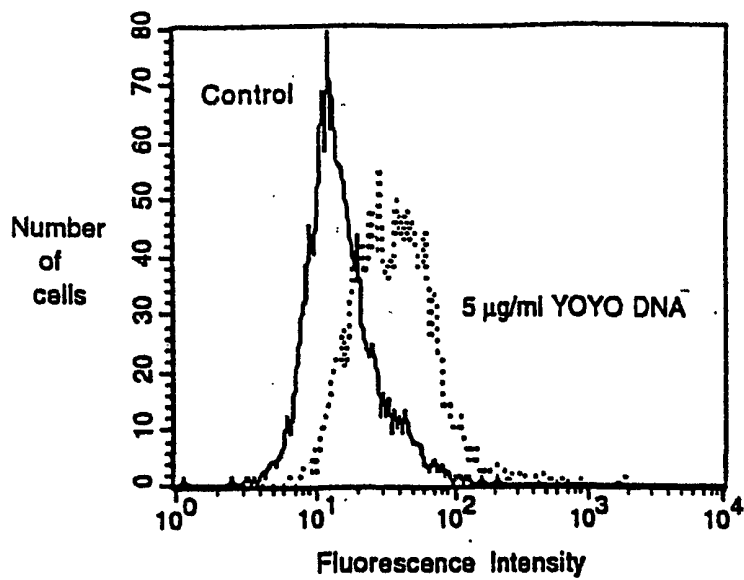


Figure 17

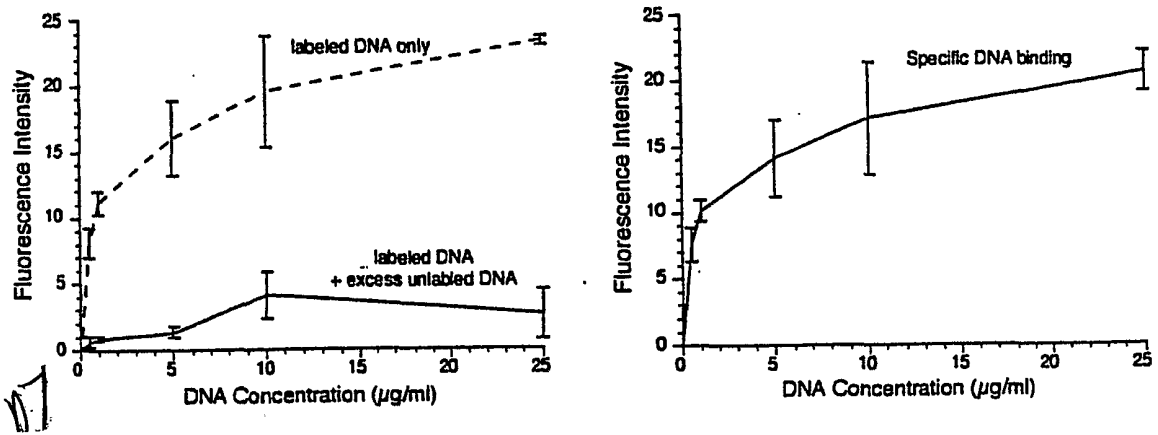


Figure 18

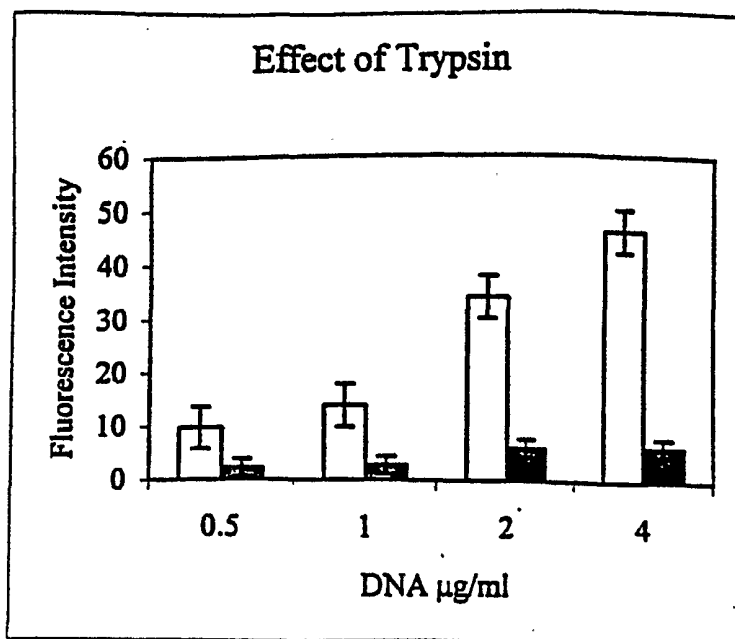


Figure 19

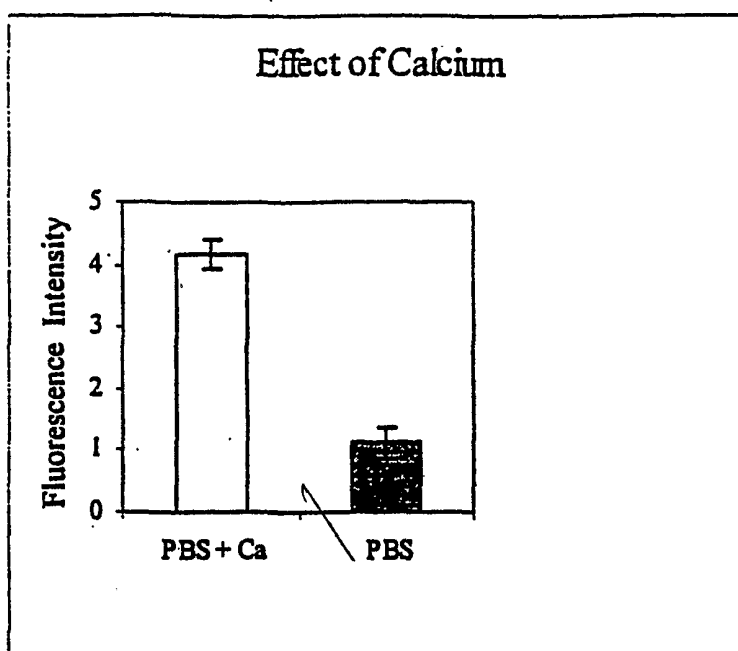


Figure 20

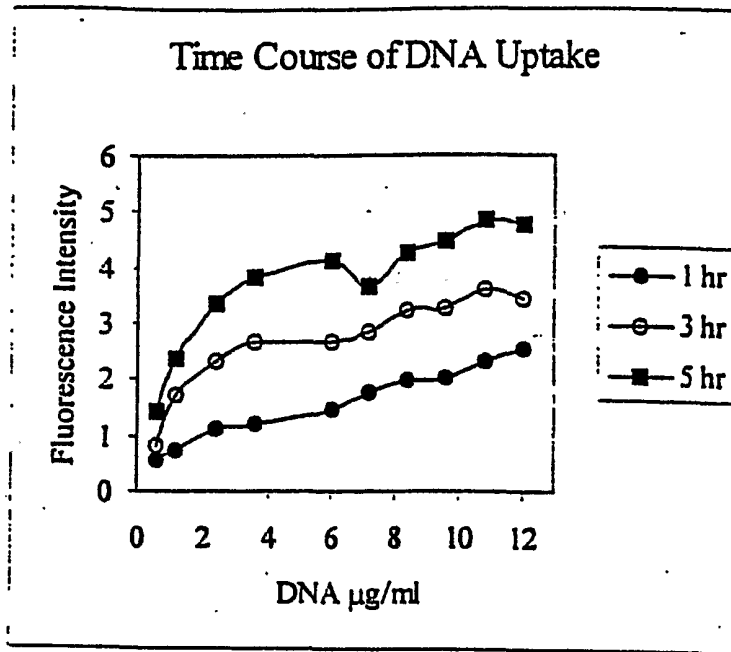


Figure 2d

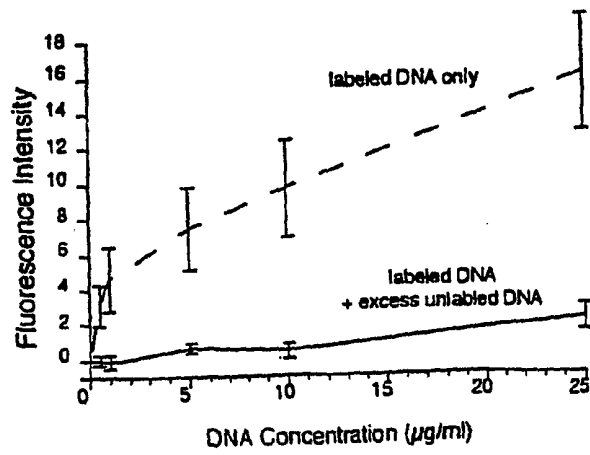


Figure 22

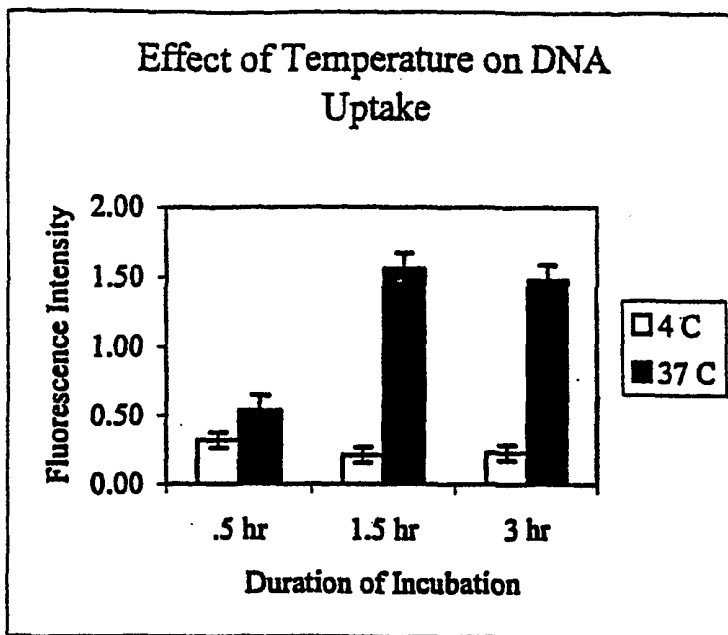
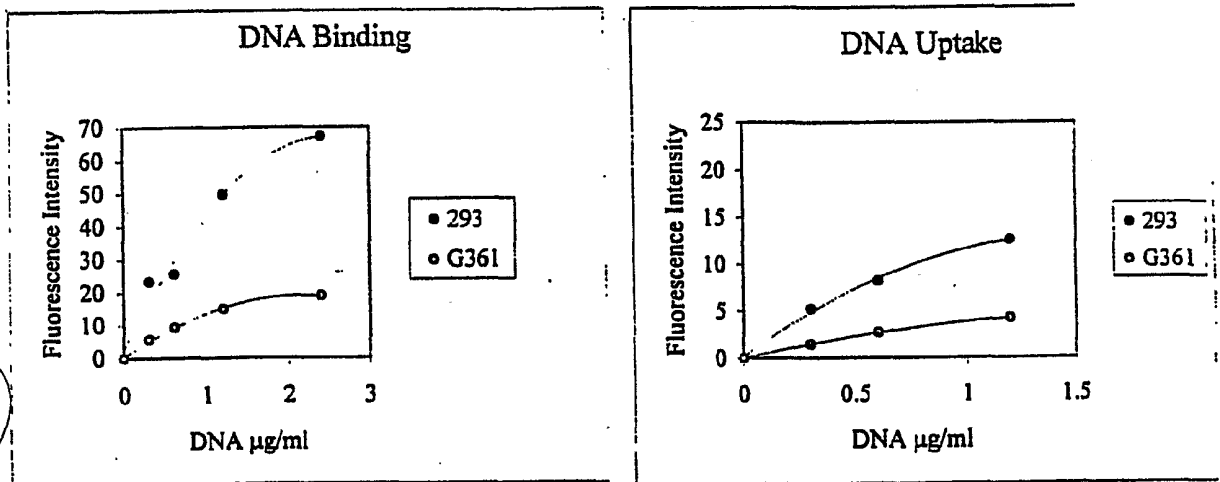


Figure 23



23

293 cells (solid circles) and G361 cells (open

Figure 2A

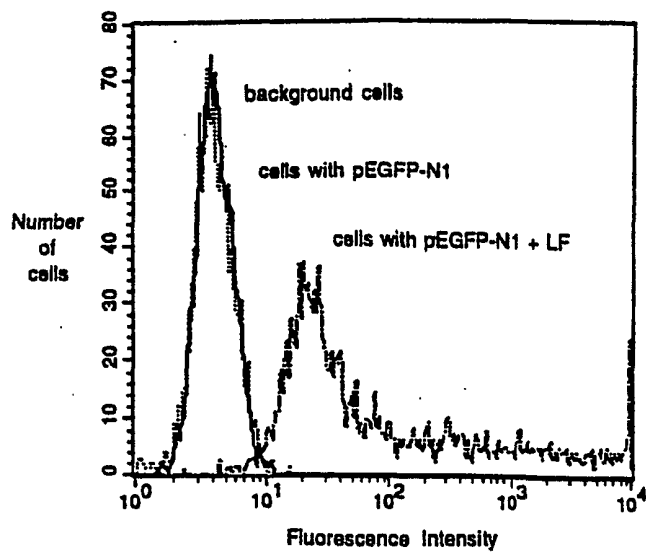
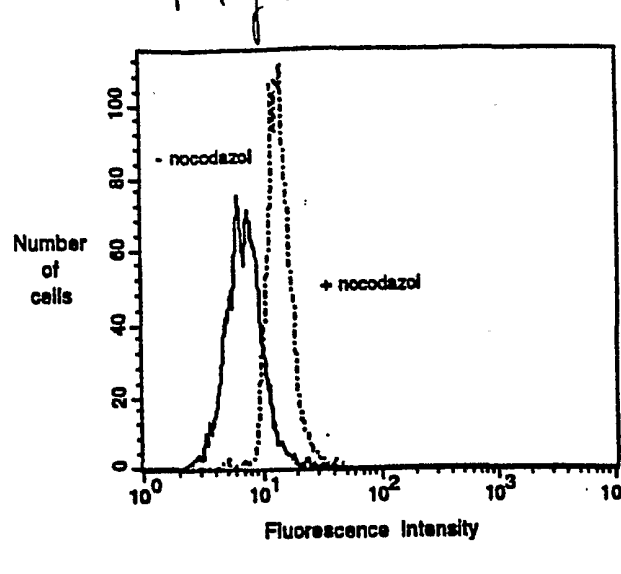


Figure 25



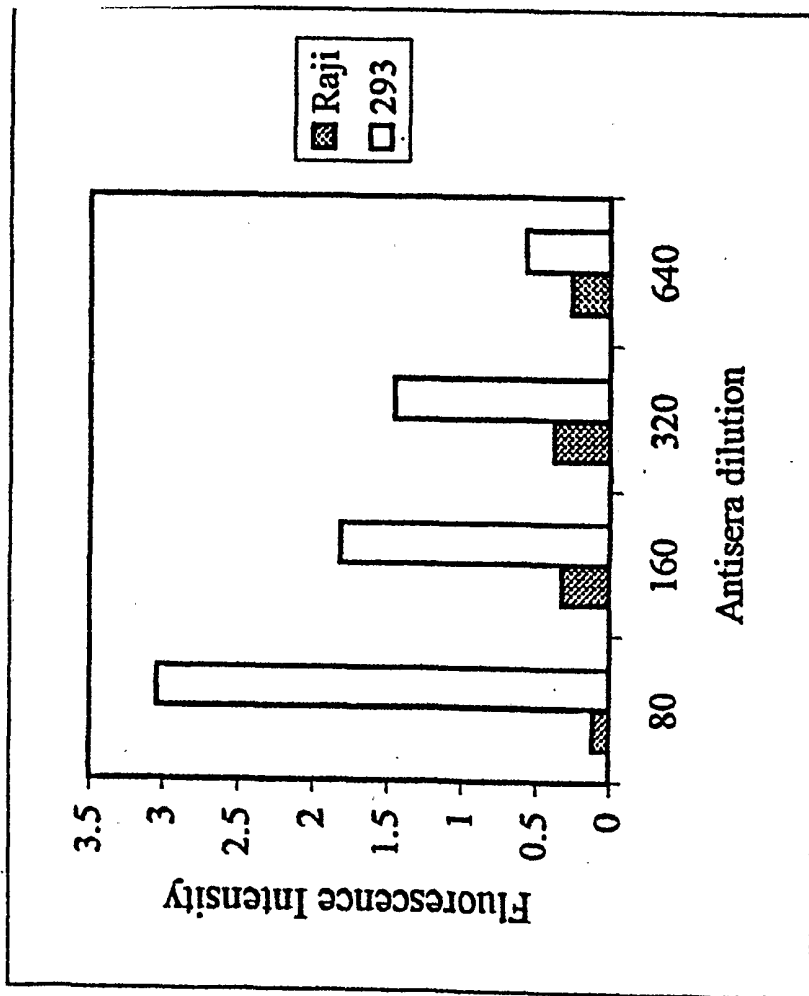


Fig 8D

SEQUENCE LISTING

<110> Hefeneider, Steven
 Merkins, Louise
 Bennett, Robert
 Seiss, Donald

<120> Mammalian DNA Binding Membrane-Associated Protein-encoding Gene and Uses

<130> 00-617-A

<160> 6

<170> PatentIn version 3.0

<210> 1

<211> 4351

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (602)..(4174)

<223> Native DNA and protein sequence

<400> 1

```

agccaagtct tgtcagagat ttcctctttc aggtggcaaa gctgttttct tcacacttga      60
gtctctacaa tattgtttgg atcagtagtt tccaaagttc attaactcct ggccatactt      120
tattatgttt tggggactg gttatccaag ggaaacactt ttttaacaaa caaaacaaaa      180
aaaccgcccc gcagtcctaaa gtaatttgtg ttctataaaa tggaatatgg aaagttaatt      240
tgcttgtttg atgtggctgt tgagaaaaat acataaaaagc tttgatgttt attatgtgag      300
caaccaatat aaatacagtt tagttgaaag gaacactatt aaggatttgt ttccaggcag      360
aatttcagaa atgtaattaa ttcagcaaat aggtttttta aaaaagacat ccaaagggta      420
taaaattatt tagaagtatt ttaggtctga agctgtaata gttgacttaa gcaattaact      480
cttcaaaggt gaatgatgaa tatgtgggta attcatactt ttgtccattt ctagcttaca      540
aaacactaca cagcaaaata atgatctgct agactgctaa cccgagcatc cagcttccac      600
a atg cct gtg cag gca gct caa tgg aca gaa ttt ctg tcc tgt cca atc      649
  Met Pro Val Gln Ala Ala Gln Trp Thr Glu Phe Leu Ser Cys Pro Ile
   1             5             10             15
tgc tat aat gaa ttt gat gag aat gtg cac aaa ccc atc agt tta ggt      697
Cys Tyr Asn Glu Phe Asp Glu Asn Val His Lys Pro Ile Ser Leu Gly
   20             25             30
tgt tca cac act gtt tgc aag acc tgc ttg aat aaa ott cat cga aaa      745
Cys Ser His Thr Val Cys Lys Thr Cys Leu Asn Lys Leu His Arg Lys
    
```

35					40					45						
gct	tgt	cct	ttt	gac	cag	act	gcc	atc	aac	aca	gat	att	gat	gta	ctt	793
Ala	Cys	Pro	Phe	Asp	Gln	Thr	Ala	Ile	Asn	Thr	Asp	Ile	Asp	Val	Leu	
50					55					60						
cct	gtc	aac	ttc	gca	ctt	ctc	cag	tta	gtt	gga	gcc	cag	gta	cca	gat	841
Pro	Val	Asn	Phe	Ala	Leu	Leu	Gln	Leu	Val	Gly	Ala	Gln	Val	Pro	Asp	
65					70					75					80	
cat	cag	tca	att	aag	tta	agt	aat	cta	ggg	gag	aat	aaa	cac	tat	gag	889
His	Gln	Ser	Ile	Lys	Leu	Ser	Asn	Leu	Gly	Glu	Asn	Lys	His	Tyr	Glu	
85					90					95						
ggt	gca	aag	aaa	tgc	ggt	gag	gat	ttg	gca	ctc	tac	tta	aaa	cca	cta	937
Val	Ala	Lys	Lys	Cys	Val	Glu	Asp	Leu	Ala	Leu	Tyr	Leu	Lys	Pro	Leu	
100					105					110						
agt	gga	ggg	aaa	ggg	gta	gct	agc	ttg	aac	cag	agt	gca	ctg	agc	cgt	985
Ser	Gly	Gly	Lys	Gly	Val	Ala	Ser	Leu	Asn	Gln	Ser	Ala	Leu	Ser	Arg	
115					120					125						
cca	atg	caa	agg	aaa	ctg	gtg	aca	ctt	gta	aac	tgt	caa	ctg	gtg	gag	1033
Pro	Met	Gln	Arg	Lys	Leu	Val	Thr	Leu	Val	Asn	Cys	Gln	Leu	Val	Glu	
130					135					140						
gaa	gaa	ggg	cgt	gta	aga	gcc	atg	cga	gca	gct	cgt	tcc	ctt	gga	gaa	1081
Glu	Glu	Gly	Arg	Val	Arg	Ala	Met	Arg	Ala	Ala	Arg	Ser	Leu	Gly	Glu	
145					150					155					160	
aga	act	gta	aca	gaa	ctg	ata	tta	cag	cac	cag	aac	cct	cag	cag	ttg	1129
Arg	Thr	Val	Thr	Glu	Leu	Ile	Leu	Gln	His	Gln	Asn	Pro	Gln	Gln	Leu	
165					170					175						
tct	gcc	aat	cta	tgg	gcc	gct	gtc	agg	gct	cga	gga	tgc	cag	ttt	tta	1177
Ser	Ala	Asn	Leu	Trp	Ala	Ala	Val	Arg	Ala	Arg	Gly	Cys	Gln	Phe	Leu	
180					185					190						
ggg	cca	gct	atg	caa	gaa	gag	gcc	ttg	aag	ctg	gtg	tta	ctg	gca	tta	1225
Gly	Pro	Ala	Met	Gln	Glu	Glu	Ala	Leu	Lys	Leu	Val	Leu	Leu	Ala	Leu	
195					200					205						
gaa	gat	ggg	tct	gcc	ctc	tca	agg	aaa	ggt	ctg	gta	ctt	ttt	ggt	gtg	1273
Glu	Asp	Gly	Ser	Ala	Leu	Ser	Arg	Lys	Val	Leu	Val	Leu	Phe	Val	Val	
210					215					220						
cag	aga	cta	gaa	cca	aga	ttt	cct	cag	gca	tca	aaa	aca	agt	att	ggg	1321
Gln	Arg	Leu	Glu	Pro	Arg	Phe	Pro	Gln	Ala	Ser	Lys	Thr	Ser	Ile	Gly	
225					230					235					240	
cat	ggt	gtg	caa	cta	ctg	tat	cga	gct	tct	tgt	ttt	aag	ggt	acc	aaa	1369
His	Val	Val	Gln	Leu	Leu	Tyr	Arg	Ala	Ser	Cys	Phe	Lys	Val	Thr	Lys	
245					250					255						
aga	gat	gaa	gac	tct	tcc	cta	atg	cag	ctg	aag	gag	gaa	ttt	cgg	agt	1417
Arg	Asp	Glu	Asp	Ser	Ser	Leu	Met	Gln	Leu	Lys	Glu	Glu	Phe	Arg	Ser	
260					265					270						

tat gaa gca tta cgc aga gaa cat gat gcc caa att gtt cat att gcc	1465
Tyr Glu Ala Leu Arg Arg Glu His Asp Ala Gln Ile Val His Ile Ala	
275 280 285	
atg gaa gca gga ctc cgt att tca cct gaa cag tgg tcc tct ctt ttg	1513
Met Glu Ala Gly Leu Arg Ile Ser Pro Glu Gln Trp Ser Ser Leu Leu	
290 295 300	
tat ggt gat ttg gct cat aaa tca cac atg cag tct atc att gat aag	1561
Tyr Gly Asp Leu Ala His Lys Ser His Met Gln Ser Ile Ile Asp Lys	
305 310 315 320	
cta cag tct cca gag tca ttt gca aag agt gtc cag gaa ttg aca att	1609
Leu Gln Ser Pro Glu Ser Phe Ala Lys Ser Val Gln Glu Leu Thr Ile	
325 330 335	
gtt ttg caa cga aca ggt gac cca gct aac tta aat aga ctg agg cct	1657
Val Leu Gln Arg Thr Gly Asp Pro Ala Asn Leu Asn Arg Leu Arg Pro	
340 345 350	
cat tta gag ctt ctt gca aac ata gac cct aat cca gac gct gtt tca	1705
His Leu Glu Leu Leu Ala Asn Ile Asp Pro Asn Pro Asp Ala Val Ser	
355 360 365	
cca act tgg gag cag ctg gaa aat gca atg gta gct gtt aaa aca gta	1753
Pro Thr Trp Glu Gln Leu Glu Asn Ala Met Val Ala Val Lys Thr Val	
370 375 380	
gtt cat ggc ctt gtg gac ttc ata caa aat tat agt aga aaa ggc cat	1801
Val His Gly Leu Val Asp Phe Ile Gln Asn Tyr Ser Arg Lys Gly His	
385 390 395 400	
gag acc cct cag cct cag cca aac agc aaa tac aag act agc atg tgc	1849
Glu Thr Pro Gln Pro Gln Pro Asn Ser Lys Tyr Lys Thr Ser Met Cys	
405 410 415	
cga gat ttg cga cag cag ggg ggt tgt cca cga gga aca aat tgt aca	1897
Arg Asp Leu Arg Gln Gln Gly Gly Cys Pro Arg Gly Thr Asn Cys Thr	
420 425 430	
ttt gcc cat tct cag gaa gag ctt gaa aag tat cga tta agg aac aaa	1945
Phe Ala His Ser Gln Glu Glu Leu Glu Lys Tyr Arg Leu Arg Asn Lys	
435 440 445	
aag atc aat gcc act gta aga acg ttt cct ctt cta aat aaa gtt ggt	1993
Lys Ile Asn Ala Thr Val Arg Thr Phe Pro Leu Leu Asn Lys Val Gly	
450 455 460	
gta aac aac act gtc aca acc aca gcc gga aat gtc att tct gtc ata	2041
Val Asn Asn Thr Val Thr Thr Thr Ala Gly Asn Val Ile Ser Val Ile	
465 470 475 480	
gga agt act gaa aca aca ggg aaa att gtt cca agt aca aac gga att	2089
Gly Ser Thr Glu Thr Thr Gly Lys Ile Val Pro Ser Thr Asn Gly Ile	
485 490 495	

tca aat gca gaa aac agt gtt tcc cag cta atc tca cgt agt act gac	2137
Ser Asn Ala Glu Asn Ser Val Ser Gln Leu Ile Ser Arg Ser Thr Asp	
500 505 510	
agt acc tta aga gct ctg gag acc gtg aag aaa gtg gga aag gtt ggc	2185
Ser Thr Leu Arg Ala Leu Glu Thr Val Lys Lys Val Gly Lys Val Gly	
515 520 525	
gct aat ggt cag aat gct gct ggg ccc tct gca gat tct gta act gaa	2233
Ala Asn Gly Gln Asn Ala Ala Gly Pro Ser Ala Asp Ser Val Thr Glu	
530 535 540	
aat aaa att ggt tct cca ccc aag act cct gta agt aat gta gca gct	2281
Asn Lys Ile Gly Ser Pro Pro Lys Thr Pro Val Ser Asn Val Ala Ala	
545 550 555 560	
acc tca gct ggg ccc tct aat gtt gga aca gag ctg aat tct gtg cct	2329
Thr Ser Ala Gly Pro Ser Asn Val Gly Thr Glu Leu Asn Ser Val Pro	
565 570 575	
caa aaa tcc agc cca ttt cta act aga gta cca gta tat cct ccg cat	2377
Gln Lys Ser Ser Pro Phe Leu Thr Arg Val Pro Val Tyr Pro Pro His	
580 585 590	
tct gaa aac att cag tat ttt caa gat cca agg act cag ata ccc ttt	2425
Ser Glu Asn Ile Gln Tyr Phe Gln Asp Pro Arg Thr Gln Ile Pro Phe	
595 600 605	
gaa gtc cca cag tac cca cag aca gga tac tat cca cca cct cca acg	2473
Glu Val Pro Gln Tyr Pro Gln Thr Gly Tyr Tyr Pro Pro Pro Pro Thr	
610 615 620	
gta cca gct ggt gtg gct ccc tgt gtt cct cgc ttt gtg agg tcc aat	2521
Val Pro Ala Gly Val Ala Pro Cys Val Pro Arg Phe Val Arg Ser Asn	
625 630 635 640	
aac gtt cca gag tcc tcc ctc cca cct gct tcc atg cca tat gcc gat	2569
Asn Val Pro Glu Ser Ser Leu Pro Pro Ala Ser Met Pro Tyr Ala Asp	
645 650 655	
cat tac agt aca ttt tcc cct cga gat cga atg aat tct tct cct tac	2617
His Tyr Ser Thr Phe Ser Pro Arg Asp Arg Met Asn Ser Ser Pro Tyr	
660 665 670	
cag cct cct cct ccg cag ccg tat gga cca gtt cct cca gta cct tct	2665
Gln Pro Pro Pro Pro Gln Pro Tyr Gly Pro Val Pro Pro Val Pro Ser	
675 680 685	
gga atg tat gct cct gtg tac gac agc agg cgc atc tgg cgc cca cct	2713
Gly Met Tyr Ala Pro Val Tyr Asp Ser Arg Arg Ile Trp Arg Pro Pro	
690 695 700	
atg tac caa cga gat gac att att aga agc aat tct tta cct cca atg	2761
Met Tyr Gln Arg Asp Asp Ile Ile Arg Ser Asn Ser Leu Pro Pro Met	
705 710 715 720	
gat gtg atg cac tca tct gtc tat cag aca tct ttg cgg gaa aga tat	2809

Asp	Val	Met	His	Ser	Ser	Val	Tyr	Gln	Thr	Ser	Leu	Arg	Glu	Arg	Tyr		
				725					730					735			
aac	tca	tta	gat	gga	tat	tat	tcg	gtg	gct	tgt	cag	cca	cca	agt	gag	2857	
Asn	Ser	Leu	Asp	Gly	Tyr	Tyr	Ser	Val	Ala	Cys	Gln	Pro	Pro	Ser	Glu		
			740					745						750			
cca	agg	aca	act	gtg	cct	tta	cca	agg	gaa	cct	tgt	ggt	cat	ttg	aag	2905	
Pro	Arg	Thr	Thr	Val	Pro	Leu	Pro	Arg	Glu	Pro	Cys	Gly	His	Leu	Lys		
		755						760						765			
acc	agt	tgc	gag	gag	cag	ata	aga	aga	aag	cca	gat	cag	tgg	gca	cag	2953	
Thr	Ser	Cys	Glu	Glu	Gln	Ile	Arg	Arg	Lys	Pro	Asp	Gln	Trp	Ala	Gln		
	770					775						780					
tac	cac	act	cag	aaa	gca	cct	ctt	gtc	tct	tca	act	ctt	cct	gtg	gca	3001	
Tyr	His	Thr	Gln	Lys	Ala	Pro	Leu	Val	Ser	Ser	Thr	Leu	Pro	Val	Ala		
785					790					795					800		
aca	cag	tca	cca	aca	cca	cct	tct	cct	ctg	ttc	agt	gta	gac	ttt	cgt	3049	
Thr	Gln	Ser	Pro	Thr	Pro	Pro	Ser	Pro	Leu	Phe	Ser	Val	Asp	Phe	Arg		
				805					810					815			
gcg	gat	ttc	tca	gag	agt	gtg	agt	ggg	aca	aaa	ttt	gaa	gaa	gat	cat	3097	
Ala	Asp	Phe	Ser	Glu	Ser	Val	Ser	Gly	Thr	Lys	Phe	Glu	Glu	Asp	His		
			820					825						830			
ctt	tcc	cat	tat	tct	ccc	tgg	tct	tgt	ggc	acc	ata	ggc	tcc	tgt	ata	3145	
Leu	Ser	His	Tyr	Ser	Pro	Trp	Ser	Cys	Gly	Thr	Ile	Gly	Ser	Cys	Ile		
		835					840					845					
aat	gcc	att	gat	tca	gag	ccc	aaa	gat	gtc	att	gct	aat	tca	aat	gct	3193	
Asn	Ala	Ile	Asp	Ser	Glu	Pro	Lys	Asp	Val	Ile	Ala	Asn	Ser	Asn	Ala		
	850					855					860						
gtg	tta	atg	gac	ctg	gac	agt	ggg	gat	ggt	aag	aga	aga	gta	cat	tta	3241	
Val	Leu	Met	Asp	Leu	Asp	Ser	Gly	Asp	Val	Lys	Arg	Arg	Val	His	Leu		
865					870					875					880		
ttt	gaa	acc	cag	aga	agg	aca	aaa	gaa	gaa	gat	cca	ata	att	ccc	ttt	3289	
Phe	Glu	Thr	Gln	Arg	Arg	Thr	Lys	Glu	Glu	Asp	Pro	Ile	Ile	Pro	Phe		
			885							890					895		
agt	gat	gga	ccc	atc	atc	tca	aaa	tgg	ggg	gcg	att	tcc	aga	tct	tcc	3337	
Ser	Asp	Gly	Pro	Ile	Ile	Ser	Lys	Trp	Gly	Ala	Ile	Ser	Arg	Ser	Ser		
			900					905						910			
cgt	aca	ggg	tac	cat	acc	aca	gat	cct	gtc	cag	gcc	act	gct	tcc	caa	3385	
Arg	Thr	Gly	Tyr	His	Thr	Thr	Asp	Pro	Val	Gln	Ala	Thr	Ala	Ser	Gln		
		915					920							925			
gga	agt	gcg	act	aag	ccc	atc	agt	gta	tca	gat	tat	gtc	cct	tat	gtc	3433	
Gly	Ser	Ala	Thr	Lys	Pro	Ile	Ser	Val	Ser	Asp	Tyr	Val	Pro	Tyr	Val		
	930					935					940						
aat	gct	ggt	gat	tca	agg	tgg	agt	tca	tat	ggc	aac	gag	gcc	aca	tca	3481	
Asn	Ala	Val	Asp	Ser	Arg	Trp	Ser	Ser	Tyr	Gly	Asn	Glu	Ala	Thr	Ser		

945	950	955	960	
tca gca cac tat gtt gaa agg gac aga ttc att gtt act gat tta tct				3529
Ser Ala His Tyr Val Glu Arg Asp Arg Phe Ile Val Thr Asp Leu Ser				
	965	970	975	
ggt cat aga aag cat tcc agt act ggg gac ctt ttg agc ctt gaa ctt				3577
Gly His Arg Lys His Ser Ser Thr Gly Asp Leu Leu Ser Leu Glu Leu				
	980	985	990	
cag cag gcc aag agc aac tca tta ctt ctt cag aga gag gcc aat gct				3625
Gln Gln Ala Lys Ser Asn Ser Leu Leu Leu Gln Arg Glu Ala Asn Ala				
	995	1000	1005	
ttg gcc atg caa cag aag tgg aat tcc ctg gat gaa ggc cgt cac				3670
Leu Ala Met Gln Gln Lys Trp Asn Ser Leu Asp Glu Gly Arg His				
	1010	1015	1020	
ctt acc tta aac ctt tta agc aag gaa att gaa cta aga aat gga				3715
Leu Thr Leu Asn Leu Leu Ser Lys Glu Ile Glu Leu Arg Asn Gly				
	1025	1030	1035	
gag tta cag agt gat tat aca gaa gat gca aca gat act aaa cct				3760
Glu Leu Gln Ser Asp Tyr Thr Glu Asp Ala Thr Asp Thr Lys Pro				
	1040	1045	1050	
gat agg gat atc gag tta gag ctt tca gca ctt gat act gat gaa				3805
Asp Arg Asp Ile Glu Leu Glu Leu Ser Ala Leu Asp Thr Asp Glu				
	1055	1060	1065	
cct gat gga caa agt gaa cca att gaa gag atc ttg gac ata cag				3850
Pro Asp Gly Gln Ser Glu Pro Ile Glu Glu Ile Leu Asp Ile Gln				
	1070	1075	1080	
ctt ggt atc agt tct caa aat gat cag ttg cta aat gga atg gca				3895
Leu Gly Ile Ser Ser Gln Asn Asp Gln Leu Leu Asn Gly Met Ala				
	1085	1090	1095	
gtg gaa aat ggg cat cca gta cag cag cac caa aag gag cca cca				3940
Val Glu Asn Gly His Pro Val Gln Gln His Gln Lys Glu Pro Pro				
	1100	1105	1110	
aag cag aag aaa cag agt tta ggt gaa gac cat gtg att ctg gag				3985
Lys Gln Lys Lys Gln Ser Leu Gly Glu Asp His Val Ile Leu Glu				
	1115	1120	1125	
gag caa aaa aca att ctg ccg gta act tct tgc ttt agc cag cca				4030
Glu Gln Lys Thr Ile Leu Pro Val Thr Ser Cys Phe Ser Gln Pro				
	1130	1135	1140	
ctc cca gtg tct att agc aat gca agt tgc ctc ccc atc acc aca				4075
Leu Pro Val Ser Ile Ser Asn Ala Ser Cys Leu Pro Ile Thr Thr				
	1145	1150	1155	
tct gtc agt gct ggc aac ctc att ctg aaa act cat gtt atg tct				4120
Ser Val Ser Ala Gly Asn Leu Ile Leu Lys Thr His Val Met Ser				
	1160	1165	1170	

gaa gat aaa aac gac ttt tta aaa cct gtt gca aat ggg aag atg 4165
 Glu Asp Lys Asn Asp Phe Leu Lys Pro Val Ala Asn Gly Lys Met
 1175 1180 1185

ggt aac agc tgaaggagg ttcattcttc aaatttgatg ccacaccatg 4214
 Val Asn Ser
 1190

gaagcattta cactagcttt ttatatatat aatatatatt atataatgta tatttttttt 4274

aaaaaaaaaga tattactggg ggcattccatt tcctgtggac tctttgatac ttcaagccct 4334

cttgcattag cattatg 4351

<210> 2
 <211> 1191
 <212> PRT
 <213> Homo sapiens

<400> 2

Met Pro Val Gln Ala Ala Gln Trp Thr Glu Phe Leu Ser Cys Pro Ile
 1 5 10 15

Cys Tyr Asn Glu Phe Asp Glu Asn Val His Lys Pro Ile Ser Leu Gly
 20 25 30

Cys Ser His Thr Val Cys Lys Thr Cys Leu Asn Lys Leu His Arg Lys
 35 40 45

Ala Cys Pro Phe Asp Gln Thr Ala Ile Asn Thr Asp Ile Asp Val Leu
 50 55 60

Pro Val Asn Phe Ala Leu Leu Gln Leu Val Gly Ala Gln Val Pro Asp
 65 70 75 80

His Gln Ser Ile Lys Leu Ser Asn Leu Gly Glu Asn Lys His Tyr Glu
 85 90 95

Val Ala Lys Lys Cys Val Glu Asp Leu Ala Leu Tyr Leu Lys Pro Leu
 100 105 110

Ser Gly Gly Lys Gly Val Ala Ser Leu Asn Gln Ser Ala Leu Ser Arg
 115 120 125

Pro Met Gln Arg Lys Leu Val Thr Leu Val Asn Cys Gln Leu Val Glu
 130 135 140

Glu Glu Gly Arg Val Arg Ala Met Arg Ala Ala Arg Ser Leu Gly Glu
 145 150 155 160
 Arg Thr Val Thr Glu Leu Ile Leu Gln His Gln Asn Pro Gln Gln Leu
 165 170 175
 Ser Ala Asn Leu Trp Ala Ala Val Arg Ala Arg Gly Cys Gln Phe Leu
 180 185 190
 Gly Pro Ala Met Gln Glu Glu Ala Leu Lys Leu Val Leu Leu Ala Leu
 195 200 205
 Glu Asp Gly Ser Ala Leu Ser Arg Lys Val Leu Val Leu Phe Val Val
 210 215 220
 Gln Arg Leu Glu Pro Arg Phe Pro Gln Ala Ser Lys Thr Ser Ile Gly
 225 230 235 240
 His Val Val Gln Leu Leu Tyr Arg Ala Ser Cys Phe Lys Val Thr Lys
 245 250 255
 Arg Asp Glu Asp Ser Ser Leu Met Gln Leu Lys Glu Glu Phe Arg Ser
 260 265 270
 Tyr Glu Ala Leu Arg Arg Glu His Asp Ala Gln Ile Val His Ile Ala
 275 280 285
 Met Glu Ala Gly Leu Arg Ile Ser Pro Glu Gln Trp Ser Ser Leu Leu
 290 295 300
 Tyr Gly Asp Leu Ala His Lys Ser His Met Gln Ser Ile Ile Asp Lys
 305 310 315 320
 Leu Gln Ser Pro Glu Ser Phe Ala Lys Ser Val Gln Glu Leu Thr Ile
 325 330 335
 Val Leu Gln Arg Thr Gly Asp Pro Ala Asn Leu Asn Arg Leu Arg Pro
 340 345 350
 His Leu Glu Leu Leu Ala Asn Ile Asp Pro Asn Pro Asp Ala Val Ser
 355 360 365

Pro Thr Trp Glu Gln Leu Glu Asn Ala Met Val Ala Val Lys Thr Val
 370 375 380

Val His Gly Leu Val Asp Phe Ile Gln Asn Tyr Ser Arg Lys Gly His
 385 390 395 400

Glu Thr Pro Gln Pro Gln Pro Asn Ser Lys Tyr Lys Thr Ser Met Cys
 405 410 415

Arg Asp Leu Arg Gln Gln Gly Gly Cys Pro Arg Gly Thr Asn Cys Thr
 420 425 430

Phe Ala His Ser Gln Glu Glu Leu Glu Lys Tyr Arg Leu Arg Asn Lys
 435 440 445

Lys Ile Asn Ala Thr Val Arg Thr Phe Pro Leu Leu Asn Lys Val Gly
 450 455 460

Val Asn Asn Thr Val Thr Thr Thr Ala Gly Asn Val Ile Ser Val Ile
 465 470 475 480

Gly Ser Thr Glu Thr Thr Gly Lys Ile Val Pro Ser Thr Asn Gly Ile
 485 490 495

Ser Asn Ala Glu Asn Ser Val Ser Gln Leu Ile Ser Arg Ser Thr Asp
 500 505 510

Ser Thr Leu Arg Ala Leu Glu Thr Val Lys Lys Val Gly Lys Val Gly
 515 520 525

Ala Asn Gly Gln Asn Ala Ala Gly Pro Ser Ala Asp Ser Val Thr Glu
 530 535 540

Asn Lys Ile Gly Ser Pro Pro Lys Thr Pro Val Ser Asn Val Ala Ala
 545 550 555 560

Thr Ser Ala Gly Pro Ser Asn Val Gly Thr Glu Leu Asn Ser Val Pro
 565 570 575

Gln Lys Ser Ser Pro Phe Leu Thr Arg Val Pro Val Tyr Pro Pro His
 580 585 590

Ser Glu Asn Ile Gln Tyr Phe Gln Asp Pro Arg Thr Gln Ile Pro Phe
595 600 605

Glu Val Pro Gln Tyr Pro Gln Thr Gly Tyr Tyr Pro Pro Pro Pro Thr
610 615 620

Val Pro Ala Gly Val Ala Pro Cys Val Pro Arg Phe Val Arg Ser Asn
625 630 635 640

Asn Val Pro Glu Ser Ser Leu Pro Pro Ala Ser Met Pro Tyr Ala Asp
645 650 655

His Tyr Ser Thr Phe Ser Pro Arg Asp Arg Met Asn Ser Ser Pro Tyr
660 665 670

Gln Pro Pro Pro Pro Gln Pro Tyr Gly Pro Val Pro Pro Val Pro Ser
675 680 685

Gly Met Tyr Ala Pro Val Tyr Asp Ser Arg Arg Ile Trp Arg Pro Pro
690 695 700

Met Tyr Gln Arg Asp Asp Ile Ile Arg Ser Asn Ser Leu Pro Pro Met
705 710 715 720

Asp Val Met His Ser Ser Val Tyr Gln Thr Ser Leu Arg Glu Arg Tyr
725 730 735

Asn Ser Leu Asp Gly Tyr Tyr Ser Val Ala Cys Gln Pro Pro Ser Glu
740 745 750

Pro Arg Thr Thr Val Pro Leu Pro Arg Glu Pro Cys Gly His Leu Lys
755 760 765

Thr Ser Cys Glu Glu Gln Ile Arg Arg Lys Pro Asp Gln Trp Ala Gln
770 775 780

Tyr His Thr Gln Lys Ala Pro Leu Val Ser Ser Thr Leu Pro Val Ala
785 790 795 800

Thr Gln Ser Pro Thr Pro Pro Ser Pro Leu Phe Ser Val Asp Phe Arg
805 810 815

Ala Asp Phe Ser Glu Ser Val Ser Gly Thr Lys Phe Glu Glu Asp His

Asp Arg Asp Ile Glu Leu Glu Leu Ser Ala Leu Asp Thr Asp Glu
 1055 1060 1065

Pro Asp Gly Gln Ser Glu Pro Ile Glu Glu Ile Leu Asp Ile Gln
 1070 1075 1080

Leu Gly Ile Ser Ser Gln Asn Asp Gln Leu Leu Asn Gly Met Ala
 1085 1090 1095

Val Glu Asn Gly His Pro Val Gln Gln His Gln Lys Glu Pro Pro
 1100 1105 1110

Lys Gln Lys Lys Gln Ser Leu Gly Glu Asp His Val Ile Leu Glu
 1115 1120 1125

Glu Gln Lys Thr Ile Leu Pro Val Thr Ser Cys Phe Ser Gln Pro
 1130 1135 1140

Leu Pro Val Ser Ile Ser Asn Ala Ser Cys Leu Pro Ile Thr Thr
 1145 1150 1155

Ser Val Ser Ala Gly Asn Leu Ile Leu Lys Thr His Val Met Ser
 1160 1165 1170

Glu Asp Lys Asn Asp Phe Leu Lys Pro Val Ala Asn Gly Lys Met
 1175 1180 1185

Val Asn Ser
 1190

<210> 3

<211> 4351

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (602)..(4174)

<223> DNA with C to G substitution changing second codon to alanine

<400> 3

agccaagtct tgtcagagat ttctctttc aggtggcaaa gctgttttct tcacacttga 60

gtctctacaa tattggttgg atcagtagtt tccaaagttc attaactcct ggccatactt 120

aga act gta aca gaa ctg ata tta cag cac cag aac cct cag cag ttg	1129
Arg Thr Val Thr Glu Leu Ile Leu Gln His Gln Asn Pro Gln Gln Leu	
165 170 175	
tct gcc aat cta tgg gcc gct gtc agg gct cga gga tgc cag ttt tta	1177
Ser Ala Asn Leu Trp Ala Ala Val Arg Ala Arg Gly Cys Gln Phe Leu	
180 185 190	
ggg cca gct atg caa gaa gag gcc ttg aag ctg gtg tta ctg gca tta	1225
Gly Pro Ala Met Gln Glu Glu Ala Leu Lys Leu Val Leu Leu Ala Leu	
195 200 205	
gaa gat ggt tct gcc ctc tca agg aaa gtt ctg gta ctt ttt gtt gtg	1273
Glu Asp Gly Ser Ala Leu Ser Arg Lys Val Leu Val Leu Phe Val Val	
210 215 220	
cag aga cta gaa cca aga ttt cct cag gca tca aaa aca agt att ggt	1321
Gln Arg Leu Glu Pro Arg Phe Pro Gln Ala Ser Lys Thr Ser Ile Gly	
225 230 235 240	
cat gtt gtg caa cta ctg tat cga gct tct tgt ttt aag gtt acc aaa	1369
His Val Val Gln Leu Leu Tyr Arg Ala Ser Cys Phe Lys Val Thr Lys	
245 250 255	
aga gat gaa gac tct tcc cta atg cag ctg aag gag gaa ttt cgg agt	1417
Arg Asp Glu Asp Ser Ser Leu Met Gln Leu Lys Glu Glu Phe Arg Ser	
260 265 270	
tat gaa gca tta cgc aga gaa cat gat gcc caa att gtt cat att gcc	1465
Tyr Glu Ala Leu Arg Arg Glu His Asp Ala Gln Ile Val His Ile Ala	
275 280 285	
atg gaa gca gga ctc cgt att tca cct gaa cag tgg tcc tct ctt ttg	1513
Met Glu Ala Gly Leu Arg Ile Ser Pro Glu Gln Trp Ser Ser Leu Leu	
290 295 300	
tat ggt gat ttg gct cat aaa tca cac atg cag tct atc att gat aag	1561
Tyr Gly Asp Leu Ala His Lys Ser His Met Gln Ser Ile Ile Asp Lys	
305 310 315 320	
cta cag tct cca gag tca ttt gca aag agt gtc cag gaa ttg aca att	1609
Leu Gln Ser Pro Glu Ser Phe Ala Lys Ser Val Gln Glu Leu Thr Ile	
325 330 335	
gtt ttg caa cga aca ggt gac cca gct aac tta aat aga ctg agg cct	1657
Val Leu Gln Arg Thr Gly Asp Pro Ala Asn Leu Asn Arg Leu Arg Pro	
340 345 350	
cat tta gag ctt ctt gca aac ata gac cct aat cca gac gct gtt tca	1705
His Leu Glu Leu Leu Ala Asn Ile Asp Pro Asn Pro Asp Ala Val Ser	
355 360 365	
cca act tgg gag cag ctg gaa aat gca atg gta gct gtt aaa aca gta	1753
Pro Thr Trp Glu Gln Leu Glu Asn Ala Met Val Ala Val Lys Thr Val	
370 375 380	
gtt cat ggc ctt gtg gac ttc ata caa aat tat agt aga aaa ggc cat	1801

Val 385	His	Gly	Leu	Val	Asp 390	Phe	Ile	Gln	Asn	Tyr 395	Ser	Arg	Lys	Gly	His 400	
gag acc cct cag cct cag cca aac agc aaa tac aag act agc atg tgc 1849																
Glu Thr Pro Gln Pro Gln Pro Asn Ser Lys Tyr Lys Thr Ser Met Cys 415																
405 410 415																
cga gat ttg cga cag cag ggg ggt tgt cca cga gga aca aat tgt aca 1897																
Arg Asp Leu Arg Gln Gln Gly Gly Cys Pro Arg Gly Thr Asn Cys Thr 430																
420 425 430																
ttt gcc cat tct cag gaa gag ctt gaa aag tat cga tta agg aac aaa 1945																
Phe Ala His Ser Gln Glu Glu Leu Glu Lys Tyr Arg Leu Arg Asn Lys 445																
435 440 445																
aag atc aat gcc act gta aga acg ttt cct ctt cta aat aaa gtt ggt 1993																
Lys Ile Asn Ala Thr Val Arg Thr Phe Pro Leu Leu Asn Lys Val Gly 460																
450 455 460																
gta aac aac act gtc aca acc aca gcc gga aat gtc att tct gtc ata 2041																
Val Asn Asn Thr Val Thr Thr Thr Ala Gly Asn Val Ile Ser Val Ile 480																
465 470 475 480																
gga agt act gaa aca aca ggg aaa att gtt cca agt aca aac gga att 2089																
Gly Ser Thr Glu Thr Thr Gly Lys Ile Val Pro Ser Thr Asn Gly Ile 495																
485 490 495																
tca aat gca gaa aac agt gtt tcc cag cta atc tca cgt agt act gac 2137																
Ser Asn Ala Glu Asn Ser Val Ser Gln Leu Ile Ser Arg Ser Thr Asp 510																
500 505 510																
agt acc tta aga gct ctg gag acc gtg aag aaa gtg gga aag gtt ggc 2185																
Ser Thr Leu Arg Ala Leu Glu Thr Val Lys Lys Val Gly Lys Val Gly 525																
515 520 525																
gct aat ggt cag aat gct gct ggg ccc tct gca gat tct gta act gaa 2233																
Ala Asn Gly Gln Asn Ala Ala Gly Pro Ser Ala Asp Ser Val Thr Glu 540																
530 535 540																
aat aaa att ggt tct cca ccc aag act cct gta agt aat gta gca gct 2281																
Asn Lys Ile Gly Ser Pro Pro Lys Thr Pro Val Ser Asn Val Ala Ala 560																
545 550 555 560																
acc tca gct ggg ccc tct aat gtt gga aca gag ctg aat tct gtg cct 2329																
Thr Ser Ala Gly Pro Ser Asn Val Gly Thr Glu Leu Asn Ser Val Pro 575																
565 570 575																
caa aaa tcc agc cca ttt cta act aga gta cca gta tat cct ccg cat 2377																
Gln Lys Ser Ser Pro Phe Leu Thr Arg Val Pro Val Tyr Pro Pro His 590																
580 585 590																
tct gaa aac att cag tat ttt caa gat cca agg act cag ata ccc ttt 2425																
Ser Glu Asn Ile Gln Tyr Phe Gln Asp Pro Arg Thr Gln Ile Pro Phe 605																
595 600 605																
gaa gtc cca cag tac cca cag aca gga tac tat cca cca cct cca acg 2473																
Glu Val Pro Gln Tyr Pro Gln Thr Gly Tyr Tyr Pro Pro Pro Pro Thr																

610	615	620	
gta cca gct ggt gtg gct ccc tgt gtt cct cgc ttt gtg agg tcc aat			2521
Val Pro Ala Gly Val Ala Pro Cys Val Pro Arg Phe Val Arg Ser Asn			
625	630	635	640
aac gtt cca gag tcc tcc ctc cca cct gct tcc atg cca tat gcc gat			2569
Asn Val Pro Glu Ser Ser Leu Pro Pro Ala Ser Met Pro Tyr Ala Asp			
	645	650	655
cat tac agt aca ttt tcc cct cga gat cga atg aat tct tct cct tac			2617
His Tyr Ser Thr Phe Ser Pro Arg Asp Arg Met Asn Ser Ser Pro Tyr			
	660	665	670
cag cct cct cct ccg cag ccg tat gga cca gtt cct cca gta cct tct			2665
Gln Pro Pro Pro Pro Gln Pro Tyr Gly Pro Val Pro Pro Val Pro Ser			
	675	680	685
gga atg tat gct cct gtg tac gac agc agg cgc atc tgg cgc cca cct			2713
Gly Met Tyr Ala Pro Val Tyr Asp Ser Arg Arg Ile Trp Arg Pro Pro			
	690	695	700
atg tac caa cga gat gac att att aga agc aat tct tta cct cca atg			2761
Met Tyr Gln Arg Asp Asp Ile Ile Arg Ser Asn Ser Leu Pro Pro Met			
	705	710	715
gat gtg atg cac tca tct gtc tat cag aca tct ttg cgg gaa aga tat			2809
Asp Val Met His Ser Ser Val Tyr Gln Thr Ser Leu Arg Glu Arg Tyr			
	725	730	735
aac tca tta gat gga tat tat tcg gtg gct tgt cag cca cca agt gag			2857
Asn Ser Leu Asp Gly Tyr Tyr Ser Val Ala Cys Gln Pro Pro Ser Glu			
	740	745	750
cca agg aca act gtg cct tta cca agg gaa cct tgt ggt cat ttg aag			2905
Pro Arg Thr Thr Val Pro Leu Pro Arg Glu Pro Cys Gly His Leu Lys			
	755	760	765
acc agt tgc gag gag cag ata aga aga aag cca gat cag tgg gca cag			2953
Thr Ser Cys Glu Glu Gln Ile Arg Arg Lys Pro Asp Gln Trp Ala Gln			
	770	775	780
tac cac act cag aaa gca cct ctt gtc tct tca act ctt cct gtg gca			3001
Tyr His Thr Gln Lys Ala Pro Leu Val Ser Ser Thr Leu Pro Val Ala			
	785	790	800
aca cag tca cca aca cca cct tct cct ctg ttc agt gta gac ttt cgt			3049
Thr Gln Ser Pro Thr Pro Pro Ser Pro Leu Phe Ser Val Asp Phe Arg			
	805	810	815
gcg gat ttc tca gag agt gtg agt ggt aca aaa ttt gaa gaa gat cat			3097
Ala Asp Phe Ser Glu Ser Val Ser Gly Thr Lys Phe Glu Glu Asp His			
	820	825	830
ctt tcc cat tat tct ccc tgg tct tgt ggc acc ata ggc tcc tgt ata			3145
Leu Ser His Tyr Ser Pro Trp Ser Cys Gly Thr Ile Gly Ser Cys Ile			
	835	840	845

aat gcc att gat tca gag ccc aaa gat gtc att gct aat tca aat gct	3193
Asn Ala Ile Asp Ser Glu Pro Lys Asp Val Ile Ala Asn Ser Asn Ala	
850 855 860	
gtg tta atg gac ctg gac agt ggt gat gtt aag aga aga gta cat tta	3241
Val Leu Met Asp Leu Asp Ser Gly Asp Val Lys Arg Arg Val His Leu	
865 870 875 880	
ttt gaa acc cag aga agg aca aaa gaa gaa gat cca ata att ccc ttt	3289
Phe Glu Thr Gln Arg Arg Thr Lys Glu Glu Asp Pro Ile Ile Pro Phe	
885 890 895	
agt gat gga ccc atc atc tca aaa tgg ggt gcg att tcc aga tct tcc	3337
Ser Asp Gly Pro Ile Ile Ser Lys Trp Gly Ala Ile Ser Arg Ser Ser	
900 905 910	
cgt aca ggt tac cat acc aca gat cct gtc cag gcc act gct tcc caa	3385
Arg Thr Gly Tyr His Thr Thr Asp Pro Val Gln Ala Thr Ala Ser Gln	
915 920 925	
gga agt gcg act aag ccc atc agt gta tca gat tat gtc cct tat gtc	3433
Gly Ser Ala Thr Lys Pro Ile Ser Val Ser Asp Tyr Val Pro Tyr Val	
930 935 940	
aat gct gtt gat tca agg tgg agt tca tat ggc aac gag gcc aca tca	3481
Asn Ala Val Asp Ser Arg Trp Ser Ser Tyr Gly Asn Glu Ala Thr Ser	
945 950 955 960	
tca gca cac tat gtt gaa agg gac aga ttc att gtt act gat tta tct	3529
Ser Ala His Tyr Val Glu Arg Asp Arg Phe Ile Val Thr Asp Leu Ser	
965 970 975	
ggt cat aga aag cat tcc agt act ggg gac ctt ttg agc ctt gaa ctt	3577
Gly His Arg Lys His Ser Ser Thr Gly Asp Leu Leu Ser Leu Glu Leu	
980 985 990	
cag cag gcc aag agc aac tca tta ctt ctt cag aga gag gcc aat gct	3625
Gln Gln Ala Lys Ser Asn Ser Leu Leu Leu Gln Arg Glu Ala Asn Ala	
995 1000 1005	
ttg gcc atg caa cag aag tgg aat tcc ctg gat gaa ggc cgt cac	3670
Leu Ala Met Gln Gln Lys Trp Asn Ser Leu Asp Glu Gly Arg His	
1010 1015 1020	
ctt acc tta aac ctt tta agc aag gaa att gaa cta aga aat gga	3715
Leu Thr Leu Asn Leu Leu Ser Lys Glu Ile Glu Leu Arg Asn Gly	
1025 1030 1035	
gag tta cag agt gat tat aca gaa gat gca aca gat act aaa cct	3760
Glu Leu Gln Ser Asp Tyr Thr Glu Asp Ala Thr Asp Thr Lys Pro	
1040 1045 1050	
gat agg gat atc gag tta gag ctt tca gca ctt gat act gat gaa	3805
Asp Arg Asp Ile Glu Leu Glu Leu Ser Ala Leu Asp Thr Asp Glu	
1055 1060 1065	

cct gat gga caa agt gaa cca att gaa gag atc ttg gac ata cag 3850
 Pro Asp Gly Gln Ser Glu Pro Ile Glu Glu Ile Leu Asp Ile Gln
 1070 1075 1080

ctt ggt atc agt tct caa aat gat cag ttg cta aat gga atg gca 3895
 Leu Gly Ile Ser Ser Gln Asn Asp Gln Leu Leu Asn Gly Met Ala
 1085 1090 1095

gtg gaa aat ggg cat cca gta cag cag cac caa aag gag cca cca 3940
 Val Glu Asn Gly His Pro Val Gln Gln His Gln Lys Glu Pro Pro
 1100 1105 1110

aag cag aag aaa cag agt tta ggt gaa gac cat gtg att ctg gag 3985
 Lys Gln Lys Lys Gln Ser Leu Gly Glu Asp His Val Ile Leu Glu
 1115 1120 1125

gag caa aaa aca att ctg ccg gta act tct tgc ttt agc cag cca 4030
 Glu Gln Lys Thr Ile Leu Pro Val Thr Ser Cys Phe Ser Gln Pro
 1130 1135 1140

ctc cca gtg tct att agc aat gca agt tgc ctc ccc atc acc aca 4075
 Leu Pro Val Ser Ile Ser Asn Ala Ser Cys Leu Pro Ile Thr Thr
 1145 1150 1155

tct gtc agt gct ggc aac ctc att ctg aaa act cat gtt atg tct 4120
 Ser Val Ser Ala Gly Asn Leu Ile Leu Lys Thr His Val Met Ser
 1160 1165 1170

gaa gat aaa aac gac ttt tta aaa cct gtt gca aat ggg aag atg 4165
 Glu Asp Lys Asn Asp Phe Leu Lys Pro Val Ala Asn Gly Lys Met
 1175 1180 1185

gtt aac agc tgaaggagg ttcattcttc aaatttgatga ccacaccatg 4214
 Val Asn Ser
 1190

gaagcattta cactagcttt ttatatatat aatatatatt atataatgta tatttttttt 4274

aaaaaaaaaga tattactggg ggcattccatt tcctgtggac tctttgatac ttcaagcct 4334

cttgcattag cattatg 4351

<210> 4
 <211> 1191
 <212> PRT
 <213> Homo sapiens

<400> 4

Met Ala Val Gln Ala Ala Gln Trp Thr Glu Phe Leu Ser Cys Pro Ile
 1 5 10 15

Cys Tyr Asn Glu Phe Asp Glu Asn Val His Lys Pro Ile Ser Leu Gly
 20 25 30

Cys Ser His Thr Val Cys Lys Thr Cys Leu Asn Lys Leu His Arg Lys
 35 40 45
 Ala Cys Pro Phe Asp Gln Thr Ala Ile Asn Thr Asp Ile Asp Val Leu
 50 55 60
 Pro Val Asn Phe Ala Leu Leu Gln Leu Val Gly Ala Gln Val Pro Asp
 65 70 75 80
 His Gln Ser Ile Lys Leu Ser Asn Leu Gly Glu Asn Lys His Tyr Glu
 85 90 95
 Val Ala Lys Lys Cys Val Glu Asp Leu Ala Leu Tyr Leu Lys Pro Leu
 100 105 110
 Ser Gly Gly Lys Gly Val Ala Ser Leu Asn Gln Ser Ala Leu Ser Arg
 115 120 125
 Pro Met Gln Arg Lys Leu Val Thr Leu Val Asn Cys Gln Leu Val Glu
 130 135 140
 Glu Glu Gly Arg Val Arg Ala Met Arg Ala Ala Arg Ser Leu Gly Glu
 145 150 155 160
 Arg Thr Val Thr Glu Leu Ile Leu Gln His Gln Asn Pro Gln Gln Leu
 165 170 175
 Ser Ala Asn Leu Trp Ala Ala Val Arg Ala Arg Gly Cys Gln Phe Leu
 180 185 190
 Gly Pro Ala Met Gln Glu Glu Ala Leu Lys Leu Val Leu Leu Ala Leu
 195 200 205
 Glu Asp Gly Ser Ala Leu Ser Arg Lys Val Leu Val Leu Phe Val Val
 210 215 220
 Gln Arg Leu Glu Pro Arg Phe Pro Gln Ala Ser Lys Thr Ser Ile Gly
 225 230 235 240
 His Val Val Gln Leu Leu Tyr Arg Ala Ser Cys Phe Lys Val Thr Lys
 245 250 255

Arg Asp Glu Asp Ser Ser Leu Met Gln Leu Lys Glu Glu Phe Arg Ser
 260 265 270

Tyr Glu Ala Leu Arg Arg Glu His Asp Ala Gln Ile Val His Ile Ala
 275 280 285

Met Glu Ala Gly Leu Arg Ile Ser Pro Glu Gln Trp Ser Ser Leu Leu
 290 295 300

Tyr Gly Asp Leu Ala His Lys Ser His Met Gln Ser Ile Ile Asp Lys
 305 310 315 320

Leu Gln Ser Pro Glu Ser Phe Ala Lys Ser Val Gln Glu Leu Thr Ile
 325 330 335

Val Leu Gln Arg Thr Gly Asp Pro Ala Asn Leu Asn Arg Leu Arg Pro
 340 345 350

His Leu Glu Leu Leu Ala Asn Ile Asp Pro Asn Pro Asp Ala Val Ser
 355 360 365

Pro Thr Trp Glu Gln Leu Glu Asn Ala Met Val Ala Val Lys Thr Val
 370 375 380

Val His Gly Leu Val Asp Phe Ile Gln Asn Tyr Ser Arg Lys Gly His
 385 390 395 400

Glu Thr Pro Gln Pro Gln Pro Asn Ser Lys Tyr Lys Thr Ser Met Cys
 405 410 415

Arg Asp Leu Arg Gln Gln Gly Gly Cys Pro Arg Gly Thr Asn Cys Thr
 420 425 430

Phe Ala His Ser Gln Glu Glu Leu Glu Lys Tyr Arg Leu Arg Asn Lys
 435 440 445

Lys Ile Asn Ala Thr Val Arg Thr Phe Pro Leu Leu Asn Lys Val Gly
 450 455 460

Val Asn Asn Thr Val Thr Thr Thr Ala Gly Asn Val Ile Ser Val Ile
 465 470 475 480

Gly Ser Thr Glu Thr Thr Gly Lys Ile Val Pro Ser Thr Asn Gly Ile

Asp Val Met His Ser Ser Val Tyr Gln Thr Ser Leu Arg Glu Arg Tyr
 725 730 735

Asn Ser Leu Asp Gly Tyr Tyr Ser Val Ala Cys Gln Pro Pro Ser Glu
 740 745 750

Pro Arg Thr Thr Val Pro Leu Pro Arg Glu Pro Cys Gly His Leu Lys
 755 760 765

Thr Ser Cys Glu Glu Gln Ile Arg Arg Lys Pro Asp Gln Trp Ala Gln
 770 775 780

Tyr His Thr Gln Lys Ala Pro Leu Val Ser Ser Thr Leu Pro Val Ala
 785 790 795 800

Thr Gln Ser Pro Thr Pro Pro Ser Pro Leu Phe Ser Val Asp Phe Arg
 805 810 815

Ala Asp Phe Ser Glu Ser Val Ser Gly Thr Lys Phe Glu Glu Asp His
 820 825 830

Leu Ser His Tyr Ser Pro Trp Ser Cys Gly Thr Ile Gly Ser Cys Ile
 835 840 845

Asn Ala Ile Asp Ser Glu Pro Lys Asp Val Ile Ala Asn Ser Asn Ala
 850 855 860

Val Leu Met Asp Leu Asp Ser Gly Asp Val Lys Arg Arg Val His Leu
 865 870 875 880

Phe Glu Thr Gln Arg Arg Thr Lys Glu Glu Asp Pro Ile Ile Pro Phe
 885 890 895

Ser Asp Gly Pro Ile Ile Ser Lys Trp Gly Ala Ile Ser Arg Ser Ser
 900 905 910

Arg Thr Gly Tyr His Thr Thr Asp Pro Val Gln Ala Thr Ala Ser Gln
 915 920 925

Gly Ser Ala Thr Lys Pro Ile Ser Val Ser Asp Tyr Val Pro Tyr Val
 930 935 940

Asn Ala Val Asp Ser Arg Trp Ser Ser Tyr Gly Asn Glu Ala Thr Ser
 945 950 955 960

Ser Ala His Tyr Val Glu Arg Asp Arg Phe Ile Val Thr Asp Leu Ser
 965 970 975

Gly His Arg Lys His Ser Ser Thr Gly Asp Leu Leu Ser Leu Glu Leu
 980 985 990

Gln Gln Ala Lys Ser Asn Ser Leu Leu Leu Gln Arg Glu Ala Asn Ala
 995 1000 1005

Leu Ala Met Gln Gln Lys Trp Asn Ser Leu Asp Glu Gly Arg His
 1010 1015 1020

Leu Thr Leu Asn Leu Leu Ser Lys Glu Ile Glu Leu Arg Asn Gly
 1025 1030 1035

Glu Leu Gln Ser Asp Tyr Thr Glu Asp Ala Thr Asp Thr Lys Pro
 1040 1045 1050

Asp Arg Asp Ile Glu Leu Glu Leu Ser Ala Leu Asp Thr Asp Glu
 1055 1060 1065

Pro Asp Gly Gln Ser Glu Pro Ile Glu Glu Ile Leu Asp Ile Gln
 1070 1075 1080

Leu Gly Ile Ser Ser Gln Asn Asp Gln Leu Leu Asn Gly Met Ala
 1085 1090 1095

Val Glu Asn Gly His Pro Val Gln Gln His Gln Lys Glu Pro Pro
 1100 1105 1110

Lys Gln Lys Lys Gln Ser Leu Gly Glu Asp His Val Ile Leu Glu
 1115 1120 1125

Glu Gln Lys Thr Ile Leu Pro Val Thr Ser Cys Phe Ser Gln Pro
 1130 1135 1140

Leu Pro Val Ser Ile Ser Asn Ala Ser Cys Leu Pro Ile Thr Thr
 1145 1150 1155

Ser Val Ser Ala Gly Asn Leu Ile Leu Lys Thr His Val Met Ser
 1160 1165 1170

Glu Asp Lys Asn Asp Phe Leu Lys Pro Val Ala Asn Gly Lys Met
 1175 1180 1185

Val Asn Ser
 1190

<210> 5
 <211> 39
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer 1 for cloning DNA-R

<400> 5
 acccgagcat ggatcckcca cmatgsctgt gcaggcagc 39

<210> 6
 <211> 26
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer 2 for cloning DNA-R

<400> 6
 ggtatctaga tccatggtgt ggtcac 26