(54) Title: NOVEL CDNA 22360 ASSOCIATED WITH RENAL DISEASE AND OTHER DISEASE STATES

(57) Abstract: The invention relates generally to changes in gene expression associated with renal disease, with inflammation, and with cell growth arrest, and to methods of using differentially expressed genes or gene products as drug discovery agents, diagnostic tools or as targets in therapeutic intervention. Specifically, the invention relates to a novel human gene that corresponds to an mRNA which is differentially-regulated in kidney tissue from patients diagnosed with immunoglobulin A nephropathy (IgAN).
NOVEL cDNA 22360 ASSOCIATED WITH RENAL DISEASE
AND OTHER DISEASE STATES

INVENTORS: Hong-Wei Sun, William E. Munger and Ronald J. Falk

CONTINUING APPLICATION INFORMATION
This application claims priority to U.S. Provisional Application No. 60/170,348 filed December 13, 1999, which is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION
The invention relates generally to changes in gene expression associated with renal disorders, with inflammatory diseases, and with cell cycles, and to methods of using such genes or gene products as research or diagnostic tools, or as targets in therapeutic intervention. Specifically, this invention relates to a novel human gene which has been identified in kidney tissue from a patient diagnosed with immunoglobulin A nephropathy (IgAN). This gene: 1) serves as a marker for renal diseases such as glomerulonephritis, including mesangial proliferation; necrotizing crescentic glomerulonephritis (NCGN); minimal change disease and sclerosis, 2) serves as a marker for inflammatory diseases including noninfectious systemic inflammatory responses (SIRS), and 3) provides a means of identifying agents that modulate its expression in renal mesangial cells and other tissues.

BACKGROUND OF THE INVENTION
Abnormal renal function is among the most common ailments requiring intensive medical care. In addition, the incidence and prevalence rates of end-stage renal disease (ESRD) in the United States continue to increase. In 1995, the incidence rate was 262 per million population, with a point prevalence rate of 975 per million population. The exact number of individuals with abnormal renal function who have not yet progressed to ESRD is difficult to assess. The incidence rate per million population of treated end-stage renal disease (ESRD) has been increasing at similar rates in most countries that record counts of new ESRD patients per year. Data from the United States Renal Data System (USRDS) suggest an

IgA nephropathy (IgAN) is the most common type of immunologically mediated glomerulonephritis (GN) and is characterized by deposition in the glomerular mesangium of IgA together with C3, C5b-9, and properdin. In patients, the co-deposition of IgA together with IgG and/or IgM can lead to a more progressive course of disease. Fifteen to forty percent of primary glomerulonephritis in parts of Europe, Asia and Japan has been linked to IgAN and it is well accepted that IgAN can lead to ESRD.

IgAN often presents either as asymptomatic microscopic hematuria and/or proteinuria (most common in adults), or episodic gross hematuria following upper respiratory and other infections or exercise. The course of IgAN is variable, with some patients showing no decline in glomerular filtration rate (GFR) over decades and others developing the nephrotic syndrome, hypertension and renal failure.

Diabetic nephropathy is also a common cause of end-stage renal disease (ESRD) and accounts for 35% of the ESRD population in the United States. It results in considerable morbidity, mortality and expense. The average cost of managing one diabetic patient with ESRD is approximately $50,000 a year (Kobrin SM. Kidney International Supplement, Vol. 63 (Dec. 1997) pp. S144-150). Approximately 5.8 million people in the United States have been diagnosed by a physician as being diabetic, and an additional four to five million people have undiagnosed diabetes. Although the incidence of new cases of diabetes appears to be declining from a peak of 300 per 100,000 population in 1973, to 230 per 100,000 in 1981, its prevalence continues to rise due to a 19 percent decline since 1970 in deaths caused by diabetes. In 1982, 34,583 deaths were attributed to diabetes, resulting in diabetes being ranked as the seventh leading underlying cause of death. Medical and surgical complications of diabetes due to macro- and microvascular disease result in 5,800 new cases of blindness, 4,500 perinatal deaths, 40,000 lower extremity amputations and 3,000 deaths due to diabetic coma (ketotic and hyperosmolar) and at least 4,000 new cases of end-stage renal disease per year (Verh K. Verhandelingen - Koninklijke Academie Voor Geneeskunde Van Belgie, Vol. 51, no. 2 (1989) pp. 81-151).

If intervention is expected to be successful in halting or slowing down renal disease progression, means of accurately assessing the early manifestations of renal disease need to be established. One way to accurately assess the early manifestations of renal disease is to identify markers which are uniquely associated with disease progression. Likewise, the development of therapeutics to prevent or repair kidney damage relies on the identification of kidney genes responsible for kidney cell growth and function.
SUMMARY OF THE INVENTION

The present invention is based on the discovery of a new gene family that is differentially expressed in various forms of renal disease, inflammation, or in various stages of cell growth, compared with normal cells and tissues. The invention includes isolated nucleic acid molecules selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2, an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NO: 2, an isolated nucleic acid molecule that hybridizes to a nucleic acid molecule comprising SEQ ID NO: 1 under conditions of sufficient stringency to produce a clear signal, an isolated nucleic acid molecule that hybridizes to a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2 under conditions of sufficient stringency to produce a clear signal, and an isolated nucleic acid molecule that encodes a protein that exhibits at least about 35% amino acid sequence identity to SEQ ID NO: 2.

The present invention further includes the nucleic acid molecules operably linked to one or more expression control elements, including vectors comprising the isolated nucleic acid molecules. The invention further includes host cells transformed to contain the nucleic acid molecules of the invention and methods for producing a protein comprising the step of culturing a host cell transformed with a nucleic acid molecule of the invention under conditions in which the protein is expressed.

The invention further provides an isolated polypeptide selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2, an isolated polypeptide comprising a functional or antigenic fragment of at least 6 amino acids of SEQ ID NO: 2, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NO: 2, an isolated polypeptide comprising naturally occurring amino acid sequence variants of SEQ ID NO: 2, and an isolated polypeptide that exhibits at least about 35% amino acid sequence identity to SEQ ID NO: 2.

The invention further provides an isolated antibody, including a monoclonal or polyclonal antibody or antigen-binding antibody fragment, that specifically binds to a polypeptide of the invention.
The invention further provides methods of identifying an agent which modulates the expression of a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 or the related proteins described herein, comprising: exposing cells which express the nucleic acid to the agent; and determining whether the agent modulates expression of said nucleic acid, thereby identifying an agent which modulates the expression of a nucleic acid encoding the protein.

The invention further provides methods of identifying an agent which modulates the level of or at least one activity of a protein comprising the sequence of SEQ ID NO: 2 or the related proteins described herein, comprising: exposing cells which express the protein to the agent; and determining whether the agent modulates the level of or at least one activity of said protein, thereby identifying an agent which modulates the level of or at least one activity of the protein.

The invention further provides methods of identifying binding partners for a protein comprising the sequence of SEQ ID NO: 2 or a related protein described herein, comprising: exposing said protein to a potential binding partner; and determining if the potential binding partner binds to said protein, thereby identifying binding partners for the protein.

The present invention further provides methods of modulating the expression of a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 or the related proteins described herein, comprising: administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein. The invention also provides methods of modulating at least one activity of a protein comprising the sequence of SEQ ID NO: 2 or the related proteins described herein, comprising: administering an effective amount of an agent which modulates at least one activity of the protein.

The present invention further includes non-human transgenic animals modified to contain the nucleic acid molecules of the invention or mutated nucleic acid molecules such that expression of the encoded polypeptides of the invention is prevented.

The invention further provides methods of diagnosing renal or other disease states, comprising: acquiring a tissue or blood or other sample from a subject; and determining the level of expression of nucleic acid molecules of the invention or polypeptides of the invention.
BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1  Figure 1 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 as determined by quantitative PCR analysis in diseased and control kidney biopsy tissues, relative to the expression of GAPDH. "Min. Ch." refers to minimum change; "IgAN" refers to IgA nephropathy; "NCGN" refers to necrotizing crescentic glomerular nephropathy.

Figure 2  Figure 2 shows the tissue distribution of RNA encoding the protein of SEQ ID NO: 2 as analyzed by Northern blot in human heart (lane 1), brain (2), placenta (3), lung (4), liver (5), muscle (6), kidney (7) and pancreas (lane 8). M=RNA marker.

Figure 3  Figure 3 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 as determined by quantitative PCR analysis in various tissues relative to the expression of GAPDH.

Figure 4  Figure 4 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 in human peripheral blood leukocytes (PBL) as determined by quantitative PCR, relative to the expression of GAPDH. PBL samples were obtained from normal human controls or patients with IgAN or with anti-neutrophil cytoplasmic autoantibody (ANCA) - associated renal disease.

Figure 5  Figure 5 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 in human umbilical vein endothelial cells (HUVEC) and human umbilical artery endothelial cells (HUAEC) as determined by quantitative PCR, relative to the expression of GAPDH. Cells were treated with TNFα at time 0 in a time course study.

Figure 6  Figure 6 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 in the endothelial cell line EA (an immortalized hybrid between primary human endothelial cells and the tumor cell line A549) as determined by quantitative PCR relative to the expression of GAPDH. Growth arrest was induced by two separate methods: 1) culture in serum free conditions for various lengths of time; 2) high density contact inhibition.

Figure 7  Figure 7 shows the expression levels of mRNA encoding the protein of SEQ ID NO: 2 as determined by quantitative PCR relative to the expression of GAPDH.
Samples are primary human renal epithelial cells under log, semi-high or high density culture conditions leading to growth arrest, and primary human renal mesangial cells under similar conditions of growth arrest as well as cytokine treatment (IL-1 and IL-6).

**Figure 8** Figure 8 shows plots predicting various structural features and the antigenic and hydrophilic regions of the amino acid sequence of SEQ ID NO: 2.

**DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT**

I. General Description

The present invention is based in part on a new gene family that is differentially expressed in renal disease, inflammation, and in various stages of cell growth, compared with normal cells or tissues. The human gene encodes a protein of 275 amino acids.

The proteins of the invention can serve as a diagnostic agent or marker or as a target for agents that modulate the expression or activity of the proteins. For example, agents may be identified that modulate biological processes associated with renal disease, kidney transplantation and kidney regeneration.

The present invention is further based on the development of methods for isolating binding partners that bind a protein of the invention. Additionally, these proteins provide novel targets for screening synthetic small molecules and combinatorial or naturally occurring compound libraries to discover novel therapeutics to regulate kidney function.

II. Specific Embodiments

A. Protein Molecules

The present invention provides isolated proteins, allelic variants of the proteins, and conservative amino acid substitutions of the proteins. As used herein, the “protein” or “polypeptide” refers, in part, to a protein that has the human amino acid sequence depicted in SEQ ID NO: 2. The invention includes naturally occurring allelic variants and proteins that have a slightly different amino acid sequence than that specifically recited above. Allelic variants, though possessing a slightly different amino acid sequence than those recited above, will still have the same or similar biological functions associated with the 275 amino acid
protein.

As used herein, the family of proteins related to the human amino acid sequence of SEQ ID NO: 2 refers, in part, to proteins that have been isolated from organisms in addition to humans. The methods used to identify and isolate other members of the family of proteins related to the 275 amino acid protein are described below.

The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein.

The proteins of the present invention further include insertion, deletion or conservative amino acid substitution variants of SEQ ID NO: 2. As used herein, a conservative variant refers to at least one alteration in the amino acid sequence that does not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure or hydrophobic/hydrophilic properties of the protein can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

Ordinarily, the allelic variants, the conservative substitution variants, and the members of the protein family, will have an amino acid sequence having at least about 35%, 40%, 50%, 60%, 65%, 70% or 75% amino acid sequence identity with the sequence set forth in SEQ ID NO: 2, more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95% sequence identity. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity (see section B for the relevant parameter).

Fusion proteins, or N-terminal, C-terminal or internal extensions, deletions, or insertions into
the peptide sequence shall not be construed as affecting homology.

Contemplated variants further include those containing predetermined mutations by, e.g., homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, mouse, rat, porcine, bovine, ovine, equine and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

The proteins of the present invention also include molecules having a portion of the amino acid sequence disclosed in SEQ ID NO: 2 such as fragments having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues of the protein. Such fragments, also referred to as peptides or polypeptides, may contain antigenic regions, functional regions of the protein identified as regions of the amino acid sequence which correspond to known protein domains, as well as regions of pronounced hydrophilicity. The regions are all easily identifiable by using commonly available protein sequence analysis software such as MacVector™ (Oxford Molecular). See also Figure 8.

As described below, members of the family of proteins can be used: 1) to identify agents which modulate at least one activity of the protein; 2) to identify binding partners for the protein, 3) as an antigen to raise polyclonal or monoclonal antibodies, 4) as a therapeutic agent or target, and 5) as a diagnostic agent or marker.

**B. Nucleic Acid Molecules**

The present invention further provides nucleic acid molecules that encode the protein having SEQ ID NO: 2 and the related proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a protein or peptide as defined above, is complementary to a nucleic acid sequence encoding such peptides, hybridizes to the complement of a nucleic acid that encodes the amino acid sequence of SEQ ID NO: 2 and remains stably bound to it under appropriate stringency conditions, encodes a polypeptide
sharing at least about 35%, 40%, 50%, 60%, 65%, 70% or 75%, preferably at least about 80%, more preferably at least about 90%, and most preferably at least about 95% or more amino acid sequence identity with the peptide sequences, or exhibits at least about 35%, 40%, 50%, 60%, 65%, 70% or 75%, preferably at least about 80%, more preferably at least about 90%, and most preferably at least about 95% or more nucleotide sequence identity over the open reading frame of SEQ ID NO: 1.

Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbones or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and non-obvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

Homology or identity at the nucleotide or amino acid sequence level is determined by BLAST (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs blastp, blastn, blastx, tblastn and tblastx (Altschul, SF et al. (1997) Nucleic Acid Res, 25:3389-3402, and Karlin et al. (1990) PNAS 87:2264-2268, fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the BLAST program is to first consider similar segments, with and without gaps, between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul et al. (Nature Genetics, Vol. 6 (1994) pp. 119-129) which is fully incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (i.e., the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter (low complexity) are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the BLOSUM62 matrix (recommended for query sequences over 85 nucleotide bases or amino acids in length (Henikoff et al. Proceedings of the National Academy of Science USA, Vol. 89 (1992) pp. 10915-10919, fully incorporated by reference).
For blastn, the scoring matrix is set by the ratios of M (i.e., the reward score for a pair of matching residues) to N (i.e., the penalty score for mismatching residues), wherein the default values for M and N are 5 and -4, respectively. Four blastn parameters were adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every wink\textsuperscript{th} position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent Blastp parameter settings were Q=9; R=2; wink=1; and gapw=32. A Bestfit comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C, or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is hybridization in 50% formamide, 5x SSC (0.75M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5x Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C. in 0.2x SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred molecules are those that hybridize under the above conditions to the complement of SEQ ID NO: 1 and which encode a functional protein. Even more preferred hybridizing molecules are those that hybridize under the above conditions to the complement strand of the open reading frame of SEQ ID NO:1.

As used herein, a nucleic acid molecule is said to be “isolated” when the nucleic acid molecule is substantially separated from contaminant nucleic acid molecules encoding other polypeptides.

The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein coding sequence. The size of the fragment will be determined by the intended use. For
example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein. For instance, fragments which encode peptides corresponding to predicted antigenic regions may be prepared (see Figure 8).

If the fragment is to be used as a nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming (see the discussion in section H). Fragments of the encoding nucleic acid molecules of the present invention (i.e., synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci et al. (Journal of the American Chemical Society, Vol. 103 (1981) pp. 3185-3191) or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can readily employ any such label to obtain labeled variants of the nucleic acid molecules of the invention.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

C. Isolation of Other Related Nucleic Acid Molecules

As described above, the identification and characterization of the human nucleic acid
molecule having SEQ ID NO: 1 allows a skilled artisan to isolate nucleic acid molecules that encode other members of the protein family in addition to the sequences herein described.

Briefly, a skilled artisan can readily use the amino acid sequence of SEQ ID NO: 2 to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as lambda gt11 library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the enzyme.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the protein family from any mammalian organism. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. A PCR denature- anneal-extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules.

D. rDNA molecules Containing a Nucleic Acid Molecule

The present invention further provides recombinant DNA molecules (rDNAs) that contain a coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation in situ. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1989. In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, e.g., protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, i.e., a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extra-chromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as E. coli. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to
occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from BioRad Laboratories, (Richmond, CA), pPL and pKK223 available from Pharmacia (Piscataway, NJ).

Expression vectors compatible with eukaryotic cells, including viral vectors, preferably those compatible with vertebrate cells such as kidney cells, can also be used to form rDNA molecules that contain a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), the vector pCDM8 described herein, and the like eukaryotic expression vectors. Vectors may be modified to include kidney cell specific promoters if needed.

Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, i.e., the neomycin phosphotransferase (neo) gene. (Southern et al. Journal of Molecular and Applied Genetics, Vol. 1, no. 4 (1982) pp. 327-341) Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

E. Host Cells Containing an Exogenously Supplied Coding Nucleic Acid Molecule

The present invention further provides host cells transformed with a nucleic acid molecule that encodes a protein of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred eukaryotic host cells include, but are not limited to, yeast, insect and mammalian cells, preferably vertebrate cells such as those
from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include Chinese
hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells
(NIH/3T3) available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), and the like
eukaryotic tissue culture cell lines.

Any prokaryotic host can be used to express a rDNA molecule encoding a protein of the
invention. The preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with a rDNA molecule of the present invention is
accomplished by well known methods that typically depend on the type of vector used and host
system employed. With regard to transformation of prokaryotic host cells, electroporation and salt
treatment methods are typically employed, see, for example, Cohen *et al.* *Proceedings of the
National Academy of Science USA*, Vol. 69, no. 8 (1972) pp. 2110-2114; and Maniatis *et al.*
Laboratory Press, 1982). With regard to transformation of vertebrate cells with vectors containing
rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed, see, for
example, Graham *et al.* *Virology*, Vol. 52, no. 2 (1973) pp. 456-467; and Wigler *et al.* *Proceedings

Successfully transformed cells, *i.e.*, cells that contain a rDNA molecule of the present
invention, can be identified by well known techniques including the selection for a selectable
marker. For example, cells resulting from the introduction of an rDNA of the present invention can
be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their
DNA content examined for the presence of the rDNA using a method such as that described by
*Biotechnic and Histochemistry*, Vol. 3 (1985) pp. 208; or the proteins produced from the cell
assayed via an immunological method.

**F. Production of Recombinant Proteins using a rDNA Molecule**

The present invention further provides methods for producing a protein of the invention
using nucleic acid molecules herein described. In general terms, the production of a recombinant
form of a protein typically involves the following steps:
First, a nucleic acid molecule is obtained that encodes a protein of the invention, such as the nucleic acid molecule comprising, consisting essentially of, or consisting of SEQ ID NO: 1, or nucleotides 121-948 of SEQ ID NO: 1, or nucleotides 121-945 of SEQ ID NO: 1. If the encoding sequence is uninterrupted by introns, as is this open reading frame, it is directly suitable for expression in any host.

The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

**G. Methods to Identify Binding Partners**

Another embodiment of the present invention provides methods for use in isolating and identifying binding partners of proteins of the invention. In detail, a protein of the invention is mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with the protein of the invention. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with a protein of the invention are separated from the mixture. The binding partner that bound to the protein of the invention can then be removed and further analyzed. To identify and isolate a
binding partner, the entire protein, for instance a protein comprising the entire amino acid sequence of SEQ ID NO: 2 can be used. Alternatively, a fragment of the protein can be used.

As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell. The preferred source of cellular extracts will be cells derived from human kidney tissue or cells, for instance, renal biopsy tissue or a tissue culture cell model of renal disease. Alternatively, cellular extracts may be prepared from normal human kidney tissue or available cell lines, particularly kidney derived cell lines.

A variety of methods can be used to obtain an extract of a cell. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

Once an extract of a cell is prepared, the extract is mixed with the protein of the invention under conditions in which association of the protein with the binding partner can occur. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a human cell. Features such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used.

After removal of non-associated cellular constituents found in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the protein can be
attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins. Alternatively, binding partners may be identified using a Far-Western assay according to the procedures of Takayama et al. *Methods in Molecular Biology*, Vol. 69 (1997) pp. 171-184 or identified through the use of epitope tagged proteins or GST fusion proteins.

Alternatively, the nucleic acid molecules of the invention can be used in a yeast two-hybrid system. The yeast two-hybrid system has been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described.

**H. Methods to Identify Agents that Modulate the Expression of the Nucleic Acids.**

Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding a protein of the invention such as a protein having the amino acid sequence of SEQ ID NO: 2. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention, for instance a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2, if it is capable of up- or down-regulating expression of the nucleic acid in a cell compared to a control.

In one assay format, cell lines that contain reporter gene fusions between the open reading frame defined by nucleotides 121-948 or nucleotide 121-945 of SEQ ID NO: 1, and/or the 5' and/or the 3' regulatory elements, and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam et al. *Analytical Biochemistry*, Vol. 188 (1990) pp. 245-254). Cell lines containing the reporter gene fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid of the invention.
Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding a protein of the invention such as the protein having SEQ ID NO: 2. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1989.

Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook et al. (Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1989); or Ausubel et al. (Current Protocols in Molecular Biology. NY, Greene Publishing Company, 1995).

Hybridization conditions are modified using known methods, such as those described by Sambrook et al. and Ausubel et al. as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a silicon chip or a porous
glass wafer. The solid support can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such solid supports and hybridization methods are widely available, for example, those disclosed by Beattie (1995) (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up or down regulate the expression of a nucleic acid encoding the protein having the sequence of SEQ ID NO: 2 are identified.

Hybridization for qualitative and quantitative analysis of mRNAs may also be carried out by using a RNase Protection Assay (i.e., RPA, see Ma et al. Methods, Vol. 10, no. 3 (1996) pp. 273-238). Briefly, an expression vehicle comprising cDNA encoding the gene product and a phage specific DNA dependent RNA polymerase promoter (e.g., T7, T3 or SP6 RNA polymerase) is linearized at the 3' end of the cDNA molecule, downstream from the phage promoter, wherein such a linearized molecule is subsequently used as a template for synthesis of a labeled antisense transcript of the cDNA by in vitro transcription. The labeled transcript is then hybridized to a mixture of isolated RNA (i.e., total or fractionated mRNA) by incubation at 45°C overnight in a buffer comprising 80% formamide, 40 mM Pipes, pH 6.4, 0.4 M NaCl and 1 mM EDTA. The resulting hybrids are then digested in a buffer comprising 40 μg/ml ribonuclease A and 2 μg/ml ribonuclease. After deactivation and extraction of extraneous proteins, the samples are loaded onto urea/polyacrylamide gels for analysis.

In another assay format, cells or cell lines are first identified that express the gene products of the invention physiologically (e.g., see for example, Figure 1 for tissue distribution via Northern blot, however, RPAs may serve the identical purpose of expression selection). Cell and/or cell lines so identified would be expected to comprise the necessary cellular machinery such that the fidelity of modulation of the transcriptional apparatus is maintained with regard to exogenous contact of agent with appropriate surface transduction mechanisms and/or the cytosolic cascades. Further, such cells or cell lines would be transduced or transfected with an expression vehicle (e.g., a plasmid or viral vector) construct comprising an operable non-translated 5'-promoter containing end of the structural gene encoding the instant gene products fused to one or more antigenic fragments, which are peculiar to the instant gene
products, wherein said fragments are under the transcriptional control of said promoter and are expressed as polypeptides whose molecular weight can be distinguished from the naturally occurring polypeptides or may further comprise an immunologically distinct tag or other detectable marker. Such a process is well known in the art (see Sambrook et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1989).

Cells or cell lines transduced or transfected as outlined above are then contacted with agents under appropriate conditions; for example, the agent in a pharmacologically acceptable excipient is contacted with cells in an aqueous physiological buffer such as phosphate buffered saline (PBS) at physiological pH, Eagles balanced salt solution (BSS) at physiological pH, PBS or BSS comprising serum or conditioned media comprising PBS or BSS and/or serum incubated at 37°C. Said conditions may be modulated as deemed necessary by one of skill in the art. Subsequent to contacting the cells with the agent, said cells will be disrupted and the polypeptides of the lysate are fractionated such that a polypeptide fraction is pooled and contacted with an antibody to be further processed by immunological assay (e.g., ELISA, immunoprecipitation or Western blot). The pool of proteins isolated from the "agent-contacted" sample will be compared with a control sample where only the excipient is contacted with the cells and an increase or decrease in the immunologically generated signal from the "agent-contacted" sample compared to the control will be used to distinguish the effectiveness of the agent.

I. Methods to Identify Agents that Modulate the Level of or at Least One Activity of the Proteins.

Another embodiment of the present invention provides methods for identifying agents that modulate the level of or at least one activity of a protein of the invention such as the protein having the amino acid sequence of SEQ ID NO: 2. Such methods or assays may utilize any means of monitoring or detecting the desired activity.

In one format, the relative amounts of a protein of the invention between a cell population that has been exposed to the agent to be tested compared to an un-exposed control
cell population may be assayed. In this format, probes such as specific antibodies are used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

Antibody probes are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides or proteins of the invention if they are of sufficient length, or, if desired, or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as BSA, KLH, or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co. (Rockford, IL.), may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein (Nature, Vol. 256, no. 5517 (Aug. 1975) pp. 495-497) or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either in vitro or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal antibodies or the polyclonal
antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive antibody fragments, such as the Fab, Fab', of F(ab')2 fragments is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

The antibodies or fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras with multiple species origin, such as humanized antibodies.

Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of a protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial library, or a growth broth of an organism.

As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a nonrandom basis which takes into account the sequence of the target site and/or its conformation in connection with the agent's action. Agents can be rationally selected or rationally designed by using the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid sequence is identical to or a derivative of any functional consensus site.

The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. Dominant negative proteins, DNAs encoding these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be introduced into cells to affect function. "Mimic" used herein refers to the modification of a region or several regions of a peptide molecule to provide a structure chemically different from the parent peptide but topographically and functionally similar to the parent peptide (see Grant GA. in: Meyers (ed.) Molecular Biology and Biotechnology (New York, VCH Publishers, 1995), pp. 659-664). A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.
The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

Another class of agents of the present invention are antibodies immunoreactive with critical positions of proteins of the invention. Antibody agents are obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies.

**J. Uses for Agents that Modulate at Least One Activity of the Proteins.**

As provided in the Examples, the proteins and nucleic acids of the invention, such as the protein having the amino acid sequence of SEQ ID NO: 2, are expressed in various types of diseased renal tissue, as well as in other organs, tissues and cells, for example during stages of inflammation associated with infiltrating leukocytes or during cell growth arrest. Renal disease and renal disorder refer to any disease of the kidney, including but not limited to IgA nephropathy (IgAN), necrotizing crescentic glomerulonephritis (NCGN), and minimal change disease. Agents that up- or down-regulate or modulate the expression of the protein or at least one activity of the protein, such as agonists or antagonists, may be used to up- or down-regulate or modulate biological and pathologic processes associated with the protein's function and activity.

As used herein, a subject can be any mammal, so long as the mammal is in need of modulation of a pathological or biological process mediated by a protein of the invention. The term "mammal" is defined as an individual belonging to the class Mammalia. The invention is particularly useful in the treatment of human subjects.

Pathological processes refer to a category of biological processes which produce a deleterious effect. For example, expression of a protein of the invention may be associated with kidney cell growth regeneration and/or recovery from kidney disease. As used herein, an
agent is said to modulate a pathological process when the agent reduces the degree or severity of the process. For instance, kidney damage or ESRD may be prevented or disease progression modulated by the administration of agents that up- or down-regulate or modulate in some way the expression or at least one activity of a protein of the invention.

The agents of the present invention can be provided alone, or in combination with other agents that modulate a particular pathological process. For example, an agent of the present invention can be administered in combination with other known drugs or may be combined with dialysis or anti-rejection drugs used during transplantation. As used herein, two agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

The agents of the present invention can be administered via parenteral, subcutaneous, intravenous, intramuscular, intraperitoneal, transdermal, or buccal routes. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

The present invention further provides compositions containing one or more agents which modulate expression or at least one activity of a protein of the invention. While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typical dosages comprise 0.1 to 100 μg/kg body wt. The preferred dosages comprise 0.1 to 10 μg/kg body weight. The most preferred dosages comprise 0.1 to 1 μg/kg body weight.

In addition to the pharmacologically active agent, the compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles
include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran. Optionally, the suspension may also contain stabilizers. Liposomes can also be used to encapsulate the agent for delivery into the cell.

The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral or topical administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

In practicing the methods of this invention, the compounds of this invention may be used alone or in combination, or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be coadministered along with other compounds typically prescribed for these conditions according to generally accepted medical practice. The compounds of this invention can be utilized in vivo, ordinarily in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice, or in vitro.

K. Transgenic Animals

Transgenic animals containing mutant, knock-out or modified genes corresponding to the cDNA sequence of SEQ ID NO: 1, or the open reading frame encoding the polypeptide sequence of SEQ ID NO: 2, or fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues, are also included in the invention.

Transgenic animals are genetically modified animals into which recombinant, exogenous or cloned genetic material has been experimentally transferred. Such genetic material is often referred to as a "transgene". The nucleic acid sequence of the transgene, in this case a form of SEQ ID NO: 1 may be integrated either at a locus of a genome where that particular nucleic acid sequence is not otherwise normally found or at the normal locus for the transgene. The
transgene may consist of nucleic acid sequences derived from the genome of the same species or of a different species than the species of the target animal.

The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability of the transgenic animal to transfer the genetic information to offspring. If such offspring in fact possess some or all of that alteration or genetic information, then they too are transgenic animals.

The alteration or genetic information may be foreign to the species of animal to which the recipient belongs, foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.


A number of recombinant or transgenic mice have been produced, including those which express an activated oncogene sequence (U.S. Patent No. 4,736,866); express simian SV40 T-antigen (U.S. Patent No. 5,728,915); lack the expression of interferon regulatory factor 1 (IRF-1) (U.S. Patent No. 5,731,490); exhibit dopaminergic dysfunction (U.S. Patent No. 5,723,719); express at least one human gene which participates in blood pressure control (U.S. Patent No. 5,731,489); display greater similarity to the conditions existing in naturally occurring Alzheimer's disease (U.S. Patent No. 5,720,936); have a reduced capacity to mediate cellular adhesion (U.S. Patent No. 5,602,307); possess a bovine growth hormone gene (Clutter et al. *Genetics*, Vol. 143, no. 4 (1996) pp. 1753-1760); or, are capable of generating a fully human antibody response (McCarthy. *The Lancet*, Vol. 349, no. 9049 (1997) pp. 405).

While mice and rats remain the animals of choice for most transgenic experimentation, in some instances it is preferable or even necessary to use alternative animal species.

The method of introduction of nucleic acid fragments into recombination competent mammalian cells can be by any method which favors co-transformation of multiple nucleic acid molecules. Detailed procedures for producing transgenic animals are readily available to one skilled in the art, including the disclosures in U.S. Patent No. 5,489,743 and U.S. Patent No. 5,602,307.

L. Diagnostic Methods

As the genes and proteins of the invention are differentially expressed in the tissue of patients with IgAN compared with normal renal tissue, they maybe used to diagnose or monitor IgAN, reanl function, or to track disease progression. One means of diagnosing a renal disease using the nucleic acid molecules or proteins of the invention involves obtaining kidney tissue from living subjects. Obtaining tissue samples from living sources is problematic for tissues such as kidney. However, due to the nature of the treatment paradigms for kidney patients, biopsy may be necessary. If possible, renal biopsy tissue may be obtained percutaneously. Another means of diagnosing renal disease using the nucleic acid molecules or proteins of the invention involves obtaining blood samples from living subjects.

The use of molecular biological tools has become routine in forensic technology. For example, nucleic acid probes comprising all or at least part of the sequence of SEQ ID NO: 1 may be used to determine the expression of a nucleic acid molecule in forensic/pathology specimens. Further, nucleic acid assays may be carried out by any means of conducting a transcriptional profiling analysis. In addition to nucleic acid analysis, forensic methods of the invention may target the proteins of the invention, particularly a protein comprising the amino
acid sequence of SEQ ID NO: 2, to determine up- or down- regulation of the genes (Shiverick et al., Biochim Biophys Acta (1975) 393:124-33).

Methods of the invention may involve treatment of tissues with collagenases or other proteases to make the tissue amenable to cell lysis (Semenov DE et al., Biull Eksp Biol Med (1987) 104:113-6). Further, it is possible to obtain biopsy samples from different regions of the kidney for analysis.

Assays to detect nucleic acid or protein molecules of the invention may be in any available format. Typical assays for nucleic acid molecules include hybridization or PCR based formats. Typical assays for the detection of proteins, polypeptides or peptides of the invention include the use of antibody probes in any available format such as in situ binding assays, etc. See Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1988. In preferred embodiments, assays are carried-out with appropriate controls.

The above methods may also be used in other diagnostic protocols, including protocols and methods to detect disease states in other tissues or organs, for example the colon, pancreas, and spleen, particularly when inflammation associated with infiltrating leukocytes is involved.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The following working examples therefore, specifically point out preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.
EXAMPLES

Example 1

Identification of Differentially Expressed Kidney mRNA

Kidney tissue was obtained from patients with various forms and stages of diagnosed renal disease.


Synthesis of cDNA was performed as previously described by Prashar et al. in WO 97/05286 and in Prashar et al. *Proceedings of the National Academy of Science USA*, Vol. 93 (1996) pp. 659-663. Briefly, cDNA was synthesized according to the protocol described in the GIBCO/BRL kit for cDNA synthesis. The reaction mixture for first-strand synthesis included 6 μg of total RNA, and 200 ng of a mixture of 1-base anchored oligo(dT) primers with all three possible anchored bases

(ACGTAATACGACTCACTATAGGGCGAATTGGGTGACTTTTTTTTTTTTTT)n1 (SEQ ID NO: 3) wherein n1=A/C or G) along with other components for first-strand synthesis reaction except reverse transcriptase. This mixture was incubated at 65°C for 5 minutes, chilled on ice and the process repeated. Alternatively, the reaction mixture may include 10 μg of total RNA, and 2 pmol of one of the 2-base anchored oligo(dT) primers a heel such as RP5.0 (CTCTCAAGGATCTTACCCTT 18 AT (SEQ ID NO: 4)), or RP6.0 (TAATACCGGCCACATAGCAT 18 CG (SEQ ID NO: 5)), or RP9.2 (CAGGGTAGAGACGCCTACGCT 18 GA (SEQ ID NO: 6)) along with other components for first-strand synthesis reaction except reverse transcriptase. This mixture was then layered with mineral oil and incubated at 65°C for seven minutes followed by 50°C for another seven minutes. At this stage, 2 μl of Superscript® reverse transcriptase (200 units/μl; GIBCO/BRL) was added quickly and mixed, and the reaction continued for one hour at 45-50°C. Second-strand synthesis was performed at 16°C for two hours. At the end of the reaction, the cDNAs were precipitated with ethanol and the yield of cDNA was calculated. In our experiments,
~200 ng of cDNA was obtained from 10 µg of total RNA.

The adapter oligonucleotide sequences were

A1 (TAGCGTCCGGCGCACGCGGCGCAG (SEQ ID NO: 7)) and
A2 (GATCCTGGCCGCTGGGCTGGTCTGTCGCGC (SEQ ID NO: 8)). One

microgram of oligonucleotide A2 was first phosphorylated at the 5’ end using T4
polynucleotide kinase (PNK). After phosphorylation, PNK was heat-denatured, and 1 µg of
the oligonucleotide A1 was added along with 10x annealing buffer (1 M NaCl/100 mM Tris-
HCl, pH8.0/10 mM EDTA, pH8.0) in a final volume of 20 µl. This mixture was then heated at
65°C for ten minutes followed by slow cooling to room temperature for thirty minutes,
resulting in formation of the Y adapter at a final concentration of 100 ng/µl. About 20 ng of
the cDNA was digested with 4 units of Bgl II in a final volume of 10 µl for thirty minutes at
37°C. Two microliters (~4 ng of digested cDNA) of this reaction mixture was then used for
ligation to 100 ng (~50-fold) of the Y-shaped adapter in a final volume of 5 µl for sixteen
hours at 15°C. After ligation, the reaction mixture was diluted with water to a final volume of
80 µl (adapter ligated cDNA concentration, ~50 pg/µl) and heated at 65°C for ten minutes to
denature T4 DNA ligase, and 2 µl aliquots (with ~100 pg of cDNA) were used for PCR.

The following sets of primers were used for PCR amplification of the adapter
ligated 3’-end cDNAs:

TGAAGCGGAGACGTCGGTCG(T)₁₈ n1, n2 (SEQ ID NO: 9) (wherein n1, n2 =
AA, AC, AG AT CA CC CG CT GA GC GG and GT) as the 3’ primer with A1 as the 5’ primer
or alternatively

RP 5.0, RP 6.0, or RP 9.2 used as 3’ primers with primer A1.1 serving as the 5’ primer.

To detect the PCR products on the display gel, 24 pmol of oligonucleotide A1 or A1.1 was 5’-
end-labeled using 15 µl of [γ-32 P]ATP (Amersham; 3000 Ci/mmol) and PNK in a final
volume of 20 µl for thirty minutes at 37°C. After heat denaturing PNK at 65°C for twenty
minutes, the labeled oligonucleotide was diluted to a final concentration of 2 µM in 80 µl with
unlabeled oligonucleotide A1.1. The PCR mixture (20 µl) consisted of 2 µl (~100 pg) of the
template, 2 µl of 10x PCR buffer (100 mM Tris·HCl, pH 8.3/500 mM KCl), 2 µl of 15 mM
MgCl₂ to yield 1.5 mM final Mg²⁺ concentration optimum in the reaction mixture, 200 µM
dNTPs, 200 nM each 5' and 3' PCR primers, and one unit of AmpliTaq Gold®. Primers and dNTPs were added after preheating the reaction mixture containing the rest of the components at 85°C. This "hot start" PCR was done to avoid artifactual amplification arising out of arbitrary annealing of PCR primers at lower temperature during transition from room temperature to 94°C in the first PCR cycle. PCR consisted of five cycles of 94°C for thirty seconds, 55°C for two minutes, and 72°C for sixty seconds followed by 25 cycles of 94°C for thirty seconds, 60°C for two minutes, and 72°C for sixty seconds. A higher number of cycles resulted in smeary gel patterns. PCR products (2.5 μl) were analyzed on 6% polyacrylamide sequencing gel. For double or multiple digestion following adapter ligation, 13.2 μl of the ligated cDNA sample was digested with a secondary restriction enzyme(s) in a final vol of 20 μl. From this solution, 3 μl was used as template for PCR. This template vol of 3 μl carried ~100 pg of the cDNA and 10 mM MgCl₂ (from the 10x enzyme buffer), which diluted to the optimum of 1.5 mM in the final PCR volume of 20 μl. Since Mg²⁺ comes from the restriction enzyme buffer, it was not included in the reaction mixture when amplifying secondarily cut cDNA. Individual cDNA fragments corresponding to mRNA species were separated by denaturing polyacrylamide gel electrophoresis and visualized by autoradiography. Bands were extracted from the display gels as described by Liang et al. (Current Opinions in Immunology, Vol. 7 (1995) pp. 274-280), reamplified using the 5' and 3' primers, and subcloned into pCR-Script with high efficiency using the PCR-Script cloning kit from Stratagene. Plasmids were sequenced by cycle sequencing on an automated sequencer. Alternatively, bands were extracted (cored) from the display gels, PCR amplified and sequenced directly without subcloning. The expression of a band corresponding to SEQ ID NO: 1 varied depending on the form of renal disease and on the stage of disease in the kidney tissue: up-regulation was observed in mild and moderate IgAN, but decreased expression was observed in necrotizing crescentic glomerulonephritis (NCGN). See Figure 1.

Example 2
Cloning of a Full Length human cDNA Corresponding to the EST fragment
The full length cDNA corresponding to the EST fragment in Example 1 was obtained by the oligo-pulling method. Briefly, a gene-specific oligo was designed based on the sequence of the EST identified in Example 1. The oligo was labeled with biotin and used to hybridize with 2 µg of single strand plasmid DNA (cDNA recombinants) from a human kidney cDNA library following the procedures of Sambrook et al. (Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1985). The hybridized cDNAs were separated by streptavidin-conjugated beads and eluted by heating. The eluted cDNA was converted to double strand plasmid DNA and used to transform E. coli cells (DH10B) and the longest cDNA was screened. After confirmation by PCR using gene-specific primers, the cDNA clone was subjected to DNA sequencing.

The nucleotide sequence of the full-length cDNA corresponding to the differentially regulated EST band is set forth in SEQ ID NO: 1. For SEQ ID NO: 1, the cDNA comprises about 2839 base pairs with an open reading frame from nucleotide 121 to 945 (to 948, including the stop codon) encoding a 275 amino acid protein, as is set forth in SEQ ID NO: 2.

The predicted structural characteristics, hydrophilicity and antigenicity of peptide fragments within SEQ ID NO: 2 are shown in Figure 8. Homology searches on the amino acid sequence of SEQ ID NO: 2 found motifs within the protein that correspond to known functional domains with complete identity: a cAMP- and cGMP- dependent protein kinase phosphorylation site is present at amino acid position 194-197; seven casein kinase II phosphorylation sites are present starting at amino acids 8, 53, 75, 204, 234, 244 and 249; and Protein Kinase C phosphorylation sites are present at amino acid 13, 131 an 193. Six amino acid N-myristoylation motifs occur throughout the protein starting at amino acids 9, 18, 36, 49, 161 and 238. Additionally, imperfect motif matches, containing only 1 mismatch, are found at amino acid 173 (an ATP/GTP-binding site (P-loop) motif), 153 (a eukaryotic thiol (cysteine) protease active site), 131 (a nuclear transition protein1 signature sequence), 92 (an RNA-binding region RNP-1 signature), 89-112 (a C2H2 type zinc finger domain), and at amino acids 89-111 (a C-type lectin domain signature and profile).
Example 3

Northern blot and PCR Expression Analysis

The tissue distribution of RNA encoding the protein of SEQ ID NO: 2 was analyzed by Northern blot as well as PCR expression analysis of RNA isolated from various tissues. RNA was isolated from various tissues such as human heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas using standard procedures. Northern blots were prepared using a probe derived from SEQ ID NO: 1 with hybridization conditions as described by Sambrook et al. (Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, 1985). See Figure 2.

Quantitative PCR Analysis of Expression Levels

Figures 3 to 7 show the results of the quantitative PCR analysis of expression levels of mRNA corresponding to SEQ ID NO: 1 in various human tissue or blood samples. Real time PCR detection was accomplished by the use of the ABI PRISM 7700 Sequence Detection System. The 7700 measures the fluorescence intensity of the sample each cycle and is able to detect the presence of specific amplicons within the PCR reaction. Each sample was assayed for the level of GAPDH and mRNA corresponding to SEQ ID NO: 1. GAPDH detection was performed using Perkin Elmer part #402869 according to the manufacturer's directions. Primers were designed from SEQ ID NO: 1 using Primer Express, a program developed by PE to efficiently find primers and probes for specific sequences. These primers were used in conjunction with SYBR green (Molecular Probes), a nonspecific double stranded DNA dye, to measure the expression level mRNA corresponding to SEQ ID NO: 1, which was then normalized to the GAPDH level in each sample.

Differentially Expressed mRNA in Renal and non-Renal Tissues

Tissue distribution analysis (see Figure 3) shows prominent expression in the glomeruli, medulla and cortex of the kidney, and additionally in the pancreas, colon and spleen, indicating that the gene and proteins of the invention could be involved in the pathogenesis of diseases affecting these organs or other tissues, particularly when
inflammation associated with infiltrating leukocytes is involved.

**Differentially Expressed mRNA in Peripheral Blood Leukocytes**

Quantitative PCR, as described above, further shows that a band corresponding to SEQ ID NO: 1 is upregulated in the circulating leukocytes of IgAN patients. This indicates that the gene and proteins of the invention have diagnostic and prognostic applications that use a patient's diseased renal tissue and/or blood, to monitor expression in circulating (peripheral blood) leukocytes and/or in infiltrating leukocytes. See Figure 4.

**Identification of Differentially Expressed mRNA in Inflammation**

A role for the genes and proteins of the invention in inflammation is demonstrated in the functional analysis of mRNA expression in two studies: treatment with the inflammatory cytokine, TNFα, elicits a complex pattern of up and down regulation in human umbilical venous endothelial cells (HUVEC) and human umbilical arterial endothelial cells (HUAEC) (Figure 5), and similar results are observed in renal mesangial cells treated with the inflammatory cytokines IL-1 and IL-6 (Figure 7).

**Differentially Expressed mRNA at various Stages of the Cell Cycle.**

A role in cell cycle events was shown by the upregulation of a band corresponding to SEQ ID NO: 1 under growth arrest conditions in endothelial cells (Figure 6) as well as renal epithelial and mesangial cells (Figure 7). This indicates that the gene and/or proteins of the invention exert important biological effects primarily under non-proliferative conditions.

Although the present invention has been described in detail with reference to examples above, it is understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims. All cited patents, patent applications and publications referred to in this application are herein incorporated by reference in their entirety.
WHAT IS CLAIMED:

1. An isolated nucleic acid molecule selected from the group consisting of: (a) an isolated nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2; (b) an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NO: 2; (c) an isolated nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule comprising SEQ ID NO: 1 under conditions of sufficient stringency to produce a clear signal; (d) an isolated nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NO: 2 under conditions of sufficient stringency to produce a clear signal; (e) an isolated nucleic acid molecule that encodes a protein that exhibits at least about 35% amino acid sequence identity to SEQ ID NO: 2; and (f) an isolated nucleic acid molecule which hybridizes to any of the above nucleic acid molecules.

2. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises the sequence of SEQ ID NO: 1.

3. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule consists of the sequence of SEQ ID NO: 1.

4. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 121-948 of SEQ ID NO: 1 or nucleotides 121-945 of SEQ ID NO: 1.

5. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule consists of nucleotides 121-948 of SEQ ID NO: 1 or nucleotides 121-945 of SEQ ID NO: 1.

6. The isolated nucleic acid molecule of any one of claims 1-5, wherein said nucleic acid molecule is operably linked to one or more expression control elements.
7. A vector comprising an isolated nucleic acid molecule of any one of claims 1-5.

8. A host cell transformed to contain the nucleic acid molecule of any one of claims 1-5.

9. A host cell comprising a vector of claim 7.

10. A host cell of claim 9, wherein said host is selected from the group consisting of prokaryotic host cells and eukaryotic host cells.

11. A method for producing a polypeptide comprising culturing a host cell transformed with the nucleic acid molecule of any one of claims 1-5 under conditions in which the protein encoded by said nucleic acid molecule is expressed.

12. The method of claim 11, wherein said host cell is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.


14. An isolated polypeptide or protein selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2, an isolated polypeptide comprising a fragment of at least 6 amino acids of SEQ ID NO: 2, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NO: 2, an isolated polypeptide comprising naturally occurring amino acid sequence variants of SEQ ID NO: 2, and an isolated peptide exhibiting at least about 35% amino acid sequence identity with SEQ ID NO: 2.

15. An isolated antibody that binds to a polypeptide of either claim 13 or 14.

16. The antibody of claim 14 wherein said antibody is a monoclonal or a polyclonal
antibody, or an immunologically reactive fragment thereof.

17. A method of identifying an agent which modulates the expression of a nucleic acid encoding a protein of claim 14 comprising:

    exposing cells which express the nucleic acid to the agent; and

    determining whether the agent modulates expression of said nucleic acid, thereby identifying an agent which modulates the expression of a nucleic acid encoding the protein.

18. A method of identifying an agent which modulates the level of or at least one activity of a protein of claim 14 comprising:

    exposing cells which express the protein to the agent;

    determining whether the agent modulates the level of or at least one activity of said protein, thereby identifying an agent which modulates the level of or at least one activity of the protein.

19. The method of claim 18, wherein the agent modulates one activity of the protein.

20. A method of identifying binding partners for a protein of claim 14, comprising:

    exposing said protein to a potential binding partner; and

    determining if the potential binding partner binds to said protein, thereby identifying binding partners for the protein.

21. A method of modulating the expression of a nucleic acid encoding a protein of claim 14 comprising:

    administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein.

22. A method of modulating at least one activity of a protein of claim 14 comprising:

    administering an effective amount of an agent which modulates at least one activity of
the protein.

23. A non-human transgenic animal modified to contain a nucleic acid molecule of any of claims 1-5.

24. The transgenic animal of claim 23, wherein the nucleic acid molecule contains a mutation that prevents expression of the encoded protein.

25. A method of diagnosing a disease state in a subject, comprising determining the level of expression of a nucleic acid molecule of any one of claims 1-5.


27. The method of claim 25, wherein the disease state is a renal disease.

28. The method of claim 25, wherein the disease state is associated with infiltrating leukocytes.

29. The method of claim 25, wherein the disease state is an inflammatory disease.

30. The method of claim 25, wherein the disease state is an inflammatory disease associated with infiltrating leukocytes.

31. The method of claim 26, wherein the disease state is a renal disease.

32. The method of claim 26, wherein the disease state is associated with infiltrating leukocytes.
33. The method of claim 26, wherein the disease state is an inflammatory disease.

34. The method of claim 26, wherein the disease state is an inflammatory disease associated with infiltrating leukocytes.

35. The method of any one of claims 27-34 wherein the expression of said nucleic acid molecule or protein is determined in a blood sample.

36. The method of any one of claims 27-34 wherein the expression of said nucleic acid molecule or protein is determined in a tissue sample.
Figure 7

Graph showing the expression levels of different cell types compared to GAPDH. The x-axis represents different cell types: Epithelia cells (Log), Epithelia cells (semi-high density), Epithelia cells (high density), Mesangial Cells (Control), Mesangial Cells (IL-1 + IL-6), Mesangial Cells (Growth Arrest). The y-axis represents the expression levels ranging from 0.0000 to 0.0120.
SEQUENCE LISTING

<110> Sun, Hong-Wei
       Munger, William E.
       Falk, Ronald J.

<120> Novel cDNA (22360) Associated with Renal Disease and Other Disease States

<130> 44921-5023-WO

<140>

<150> US 60/170,348
<151> 1999-12-13

<160> 9

<170> PatentIn Ver. 2.1

<210> 1
<211> 2839
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (121)-(948)

<400> 1

```
gggccccggt gtcagccac agccagcagc ctctacctct cacaaggtg ttcgaggagca 60
```
```
ttcagggcg gaagacttgg ctgcctgccc tgccaggact ttgtctctac ccctggcacc 120
```
```
atg cag ctc cag gtg ttc tgg act ggg ctt gaa tac acc tgc cgg ctc 168
Met Gln Leu Gln Val Phe Trp Thr Gly Leu Glu Tyr Thr Cys Arg Leu
   1      5     10      15
```
```
ctg ggc atc acc act gct gca gtt ttc atc ggc gtt ggc act gag acc 216
Leu Gly Ile Thr Thr Ala Ala Val Leu Ile Gly Val Gly Thr Glu Thr
   20     25     30
```
```
ttc ctc cag ggc cag ttc aaa agc ctt gct ttc tat ctt gtt gct ttt tta 264
Phe Leu Gln Gly Gln Phe Lys Ser Ser Ala Phe Tyr Leu Leu Phe Thr
   35     40     45
```
```
gga gcc gct gtc tcc ata tgt gaa ggg gcc tac ttt gtt gct gct cag ctc 312
Gly Ala Ala Val Ser Ile Cys Glu Gly Ala Tyr Phe Val Ala Gln Leu
   50     55     60
```
```
ctg gcc atc tgc ttc cag tgt caa cca ggg tcc ctt gca gac aga gta 360
Leu Ala Ile Cys Phe Glu Cys Glu Pro Gly Ser Leu Ala Asp Arg Val
   65     70     75     80
```
<210> 2
<211> 275
<212> PRT
<213> Homo sapiens

<400> 2
Met Gln Leu Gln Val Phe Trp Thr Gly Leu Glu Tyr Thr Cys Arg Leu
     1    5    10    15
Leu Gly Ile Thr Thr Ala Ala Val Leu Ile Gly Val Gly Thr Glu Thr
     20   25    30
Phe Leu Gln Gly Gln Phe Lys Ser Leu Ala Phe Tyr Leu Leu Phe Thr
     35    40    45
Gly Ala Ala Val Ser Ile Cys Gly Gly Ala Tyr Phe Val Ala Gln Leu
     50    55    60
Leu Ala Ile Cys Phe Gln Cys Gln Pro Gly Ser Leu Ala Asp Arg Val
     65    70    75    80
Arg Glu Lys Ala His Trp Leu Gly Cys Phe Gln Lys Phe Leu Ala Tyr
     85    90    95
Leu Leu Leu Ser Val Ala Cys Phe Leu His Pro Val Leu Val Trp His
    100   105   110
Val Thr Ile Pro Gly Ser Met Leu Ile Ile Thr Gly Leu Ala Tyr Phe
    115   120   125
Leu Leu Ser Lys Arg Lys Lys Arg Lys Ala Ala Pro Glu Val Leu Ala
    130   135   140
Ser Pro Glu Gln Tyr Thr Asp Pro Ser Ser Ser Ala Val Ser Thr Thr
    145   150   155   160
Gly Ser Gly Asp Thr Glu Gln Thr Tyr Thr Phe His Gly Ala Leu Lys
    165   170   175
Glu Gly Pro Ser Ser Leu Phe Ile His Met Lys Ser Ile Leu Lys Gly
    180   185   190
Thr Lys Lys Pro Ser Ala Leu Gln Pro Pro Asn Thr Leu Met Glu Leu
    195   200   205
Ser Leu Glu Pro Ala Asp Ser Leu Ala Lys Lys Gln Val His Phe
    210   215   220
Glu Asp Asn Leu Val Arg Ile Val Pro Ser Leu Ala Glu Gly Leu Asp
    225   230   235   240
Asp Gly Asp Ser Glu Pro Glu Glu Thr Thr Ser Asp Thr Thr Pro Ile
    245   250   255
Ile Pro Pro Pro Gln Ala Pro Leu Phe Leu Ser Ser Leu Thr Ala Thr
    260   265   270
Gly Leu Phe
    275

<210> 3
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer for cDNA synthesis
<220>
variation
<222> (55)
<223> v = a or c or g

<400> 3
acgtaatacg actcactata gggcgaattg ggtcgaacttt tttttttttt ttttv

<210> 4
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer for cDNA synthesis

<400> 4
cctctcaagga tcttaccgct tttttttttt ttttttttat

<210> 5
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer for cDNA synthesis

<400> 5
taataccggc ccacatagca tttttttttt ttttttttcg

<210> 6
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer for cDNA synthesis

<400> 6
cagggtagac gacgctacgc tttttttttt ttttttttga

<210> 7
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Adapter sequence for cloning
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Adaptor sequence for cloning

<400> 8
gatcctggcc gtcggctgct gtcggccgc

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<220>
<221> variation
<222> (39) .. (40)
<223> v = a or c or g; n = a or c or g or t

<400> 9
tgaagccgag acgtcggtcg tttttttttttttttttttn

PCT/US00/33658
### A. CLASSIFICATION OF SUBJECT MATTER

- **IPCs**: A61K 39/395, C07H 21/04, C12P 21/06, C07K 1/00, C07K 17/00, C07K 16/00, G01N 33/00
- **US CL**: 424/130.1, 536/23.5, 435/69.1, 435/252.3, 536/24.5, 530/350, 530/387.1, 530/388.1, 800/3

According to International Patent Classification (IPC) or to both national classification and IPC.

### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)


Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched.

Electronic database consulted during the international search (name of database and, where practicable, search terms used)

Please See Continuation Sheet

### C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Database Medline, US National Library of Medicine, (Bethesda, MD, USA), No. AA 873801, STRUSBERG, R. Human Cancer Genome Anatomy Project, March 1998.</td>
<td>1-5</td>
</tr>
</tbody>
</table>

- Further documents are listed in the continuation of Box C.
- See patent family annex.

| "P" | document published prior to the international filing date but later than the priority date claimed |
| "Y" | document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) |
| "O" | document referring to an oral disclosure, use, exhibition or other means |
| "E" | earlier application or patent published on or after the international filing date of the patent application |
| "A" | document defining the general state of the art which is not considered to be of particular relevance |
| "F" | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| "X" | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| "Y" | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "&" | document member of the same patent family |

Date of the actual completion of the international search: 23 APR 2001

Date of mailing of the international search report: [Signature]

Name and mailing address of the ISA/US
- Commissioner of Patents and Trademarks
- Box PCT
- Washington, D.C. 2031

Facsimile No. (703)305-3230

Authorized officer
- Phuong N. Huynh
- Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)
Continuation of B. FIELDS SEARCHED Item3: APS, STN, Medline, Biosis, Capulus, Embase, Sciserach, SwissProt_39, PIR_66, SPTREMBL_15, Issued Patents
IgA nephropathy, Renal disease, End stage renal disease, Inflammatory disease, Transgenic, Antibodies, Monoclonal, Polyclonal