PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A61K 38/00, 48/00, C07H 21/00, C07K 14/00, C12N 15/00, C12P 21/00

A1

(11) International Publication Number: **WO 97/20573**

(43) International Publication Date: 12 June 1997 (12.06.97)

(21) International Application Number: PCT/US95/15883

(22) International Filing Date: 4 December 1995 (04.12.95)

(71) Applicants *(for all designated States except US)*: SMITHKLINE BEECHAM CORPORATION [US/US]; Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US). HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850-3338 (US). JOSLIN DIABETES CENTER, INC. [US/US]; One Joslin Place, Boston, MA 02215 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: DUNNINGTON, Damien [GB/US]; 23 Forsythia Court, Lafayette Hill, PA 19444 (US). NI, Jian [CN/US]; 9410 Key West Avenue, Rockville, MD 20850-3338 (US). SHOELSON, Steven, E. [US/US]; 8 Mocassin Path, Natick, MA 01760 (US).

(74) Agents: BAUMEISTER, Kirk et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, PT, RO, RU, SD, SG, SI, SK, TJ, TM, TT, UA, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: GROWTH FACTOR RECEPTOR-BINDING PROTEIN 2 HOMOLOG

(57) Abstract

Isolated nucleic acid encoding a growth factor receptor binding protein-2 homolog, protein obtainable from the nucleic acid, recombinant host cells transformed with the nucleic acid and use of the protein and nucleic acid sequence are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

GROWTH FACTOR RECEPTOR-BINDING PROTEIN 2 HOMOLOG

Field of the Invention

The present invention relates to an isolated human homolog of growth factor receptor-binding protein 2 (Grb2-1) gene; to essentially pure human Grb2-1 protein; and to compositions and methods of producing and using human Grb2-1 sequences and proteins.

Background of the Invention

A number of polypeptide growth factors and hormones mediate their cellular effects through a signal transduction pathway. Transduction of signals from the cell surface receptors for these ligands to intracellular effectors frequently involves phosphorylation or dephosphorylation of specific protein substrates by regulatory protein tyrosine kinases (PTK) and phosphatases. Tyrosine phosphorylation is a major mediator of signal transduction in multicellular organisms Receptor-bound, membrane-bound and intracellular PTKs regulate cell proliferation, cell differentiation and signalling processes in immune system cells.

Aberrant protein tyrosine kinase activity has been implicated or is suspected in a number of pathologies such as diabetes, atherosclerosis, psoriasis, septic shock, bone loss, anemia, many cancers and other proliferative diseases. Accordingly, tyrosine kinases and the signal transduction pathways which they are part of are potential targets for drug design. For a review, see Levitzki *et al.* in *Science 267*, 1782-1788 (1995).

Many of the proteins comprising signal transduction pathways are present at low levels and often have opposing activities. The properties of these signalling molecules allow the cell to control transduction by means of the subcellular location and juxtaposition of effectors as well as by balancing activation with

repression such that a small change in one pathway can achieve a switching effect.

The formation of transducing complexes by juxtaposition of the signalling molecules through protein-protein interactions are mediated by specific docking domain sequence motifs. Src homology 2 (SH2) domains, which are conserved non-catalytic sequences of approximately 100 amino acids found in a variety of signalling molecules such as non-receptor PTKs and kinase target effector molecules and in oncogenic proteins, play a critical role. The SH2 domains are highly specific for short phosphotyrosine-containing peptide sequences found in autophosphorylated PTK receptors or intracellular tyrosine kinases. Src homology 3 (SH3) domains, conserved sequences of approximately 50 amino acids that mediate protein-protein interactions through sequence-specific binding to proline-rich motifs in target proteins, are also critically involved in signal transduction. Approximately 60 proteins having conserved SH2 and SH3 domains, and, in many cases, distinct catalytic domains, are now known.

One approach towards the pharmacological regulation of signal transduction pathways is to design inhibitory ligands which selectively bind to a chosen SH2 domain and thus block the interaction of a phosphorylated protein tyrosine kinase with its SH2-containing target molecule, thereby disrupting signal transduction. Any selective inhibitors would provide a useful lead for drug development.

Growth factor receptor binding protein 2 (Grb2) is a cytoplasmic signalling molecule containing an SH2 domain and two SH3 domains with a wide tissue and cell distribution. The molecule was first described by Lowenstein *et al.* in *Cell 70*, 431-442 (1992) and U.S. Patent No. 5,434,064 of Schlessinger *et al.* Interaction of Grb2 with growth factor receptors such as the

epidermal growth factor receptor (EGFR) and platelet-derived growth factor receptor (PDGFR) is mediated by the SH2 domain, can be dependent upon receptor tyrosine autophosphorylation and involves a direct interaction between Grb2 and the phosphorylated receptors. Grb2 also associates with other cellular tyrosine-phosphorylated proteins such as IRS-1 and SHC, possibly through the concerted action of its SH2 and SH3 domains. The 25 kDa protein has been identified as the human homolog of the *C. elegans* gene product *sem-5*, which plays a crucial developmental role.

Further, Grb2 has been shown to link tyrosine-phosphorylated proteins to a guanine nucleotide releasing factor of the son of sevenless (Sos) class and is considered to be essential for Ras activation and endocytosis. Studies suggest that grb2 forms a complex with activated monocyte-macrophage colony stimulating factor (M-CSF) receptor and that M-CSF stimulation of myeloid cells could activate Ras Sos1 by Grb2 binding to either M-CSF receptor, src homology 2-collagen (SHC) or P150. Thus, Grb2 could regulate the growth and differentiation of progenitor cells along the monocyte/macrophage lineage.

An isoform of *grb2* cDNA having a deletion in the SH2 domain was cloned, expressed and characterized by Fath *et al.* (*Science 264*, 971-974 (1994)). The protein encoded by the cDNA, Grb3-3, had apoptotic properties and did not bind to phosphorylated EGFR but inhibited EGF-induced transactivation of a Ras-responsive element. High amounts of Grb3-3 mRNA were expressed in rat thymus at an age when massive negative selection of thymocytes occurs. The authors state that their findings indicate that Grb3-3 may trigger active programmed cell death in the thymus and other tissues by acting as a dominant negative protein over Grb2 and suppressing proliferative signals.

The involvement of Grb2 and Grb3-3 in the signal transduction of the cell division pathway necessitates the identification of other human Grb2 homologs and isoforms and their cDNAs. A need also exists for compounds which modulate the activity of Grb2 homologs and isoforms, for methods to identify such modulators and for reagents useful in such methods.

Summary of the Invention

Accordingly, one aspect of the present invention is an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding human Grb2-1 having the nucleotide sequence as set forth in SEQ ID NO:1 from nucleotide 40 to 691;

(b) a polynucleotide capable of hybridizing to the complement of a polynucleotide according to (a) under moderately stringent hybridization conditions and which encodes a functional human Grb2-1; and

(c) a degenerate polynucleotide according to (a) or (b).

Another aspect of the invention is a functional polypeptide encoded by the polynucleotides of the invention.

Another aspect of the invention is a method for preparing essentially pure human Grb2-1 protein comprising culturing a recombinant host cell comprising a vector comprising a polynucleotide of the invention under conditions promoting expression of the protein and recovery thereof.

Another aspect of the invention is an antisense oligonucleotide comprising a sequence which is capable of binding to the polynucleotide of the invention.

Another aspect of the invention is a modulator of the polypeptides of the invention.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates Grb2-1 activity comprising the steps of:

(a) providing a Grb2 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular binding partner or synthetic analog thereof;

(b) incubating with a test substance which is suspected of modulating Grb2-1 activity under conditions which permit the formation of a Grb2-1 protein/cellular binding partner complex;

(c) assaying for the presence of the complex, free Grb2-1 protein or free cellular binding partner; and

(d) comparing to a control to determine the effect of the substance.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates Grb2-1 activity by direct binding to Grb2-1 protein comprising the steps of:

(a) providing a labelled Grb2-1 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof

(b) providing solid support-associated modulator candidates;

(c) incubating a mixture of the labelled Grb2-1 protein with the support-associated modulator candidates under conditions which can permit the formation of a Grb2-1 protein/modulator candidate complex;

(d) separating the solid support from free soluble labelled Grb2-1 protein;

(e) assaying for the presence of solid support-associated labelled protein;

(f) isolating the solid support complexed with labelled Grb2-1 protein; and

(g) identifying the modulator candidate.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates Grb2-1 activity by causing Grb2-1 to become membrane bound comprising the steps of:

(a) providing a Grb2-1 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular membrane;

(b) incubating with a test substance which is suspected of modulating Grb2-1 activity under conditions which permit the formation of a Grb2-1 protein/cellular membrane complex;

(c) assaying for the presence of the membrane-bound complex, free Grb2-1 protein or free cellular membrane; and

(d) comparing to a control to determine the effect of the substance.

Another aspect of the invention is Grb2-1 protein modulating compounds identified by the methods of the invention.

Another aspect of the invention is a method for the treatment of a patient having need to modulate Grb2-1 activity comprising administering to the patient a therapeutically effective amount of the modulating compounds of the invention.

Brief Description of the Drawings

Figure 1 is an amino acid sequence alignment of human Grb2-1 with murine Grb2.

Figure 2 is an amino acid sequence alignment of human Grb2-1 with human Grb2.

Figure 3 is a Northern blot of 16 human tissues probed with radiolabelled human *grb2* and *grb2-1* cDNAs.

Detailed Description of the Invention

As used herein, the term "*grb2-1* gene" refers to DNA molecules comprising a nucleotide sequence that encodes a homolog of human growth factor receptor

binding protein 2 which exhibits T-cell specificity. The *grb2-1* gene sequence is listed in SEQ ID NO:1. The coding region of the *grb2-1* gene consists of nucleotides 40-691 of SEQ ID NO:1. The deduced 217 amino acid sequence of the *grb2-1* gene product Grb2-1 is listed in SEQ ID NO:2.

As used herein, the term "functional fragments" when used to modify a specific gene or gene product means a less than full length portion of the gene or gene product which retains substantially all of the biological function associated with the full length gene or gene product to which it relates. To determine whether a fragment of a particular gene or gene product is a functional fragment, fragments are generated by well-known nucleolytic or proteolytic techniques and the fragments tested for the described biological function.

As used herein, an "antigen" refers to a molecule containing one or more epitopes that will stimulate a host's immune system to make a humoral and/or cellular antigen-specific response. The term is also used herein interchangeably with "immunogen."

As used herein, the term "epitope" refers to the site on an antigen or hapten to which a specific antibody molecule binds. The term is also used herein interchangeably with "antigenic determinant" or "antigenic determinant site."

As used herein, "monoclonal antibody" is understood to include antibodies derived from one species (e.g., murine, rabbit, goat, rat, human, etc.) as well as antibodies derived from two (or perhaps more) species (e.g., chimeric and humanized antibodies).

As used herein, a coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one

another so long as the expressed sequence is ultimately processed to produce the desired protein.

As used herein, "recombinant" polypeptides refer to polypeptides produced by recombinant DNA techniques; i.e., produced from cells transformed by an exogenous DNA construct encoding the desired polypeptide. "Synthetic" polypeptides are those prepared by chemical synthesis.

As used herein, a "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*; i.e., capable of replication under its own control.

As used herein, a "vector" is a replicon, such as a plasmid, phage, or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

As used herein, a "reference" gene refers to the wild type human Grb2-1 gene sequence of the invention and is understood to include the various sequence polymorphisms that exist, wherein nucleotide substitutions in the gene sequence exist, but do not affect the essential function of the gene product.

As used herein, a "mutant" gene refers human Grb2-1 sequences different from the reference gene wherein nucleotide substitutions and/or deletions and/or insertions result in perturbation of the essential function of the gene product.

As used herein, a DNA "coding sequence of" or a "nucleotide sequence encoding" a particular protein, is a DNA sequence which is transcribed and translated into a polypeptide when placed under the control of appropriate regulatory sequences.

As used herein, a "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bound at its 3' terminus by a translation start codon (e.g., ATG)

of a coding sequence and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Prokaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

As used herein, DNA "control sequences" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers and the like, which collectively provide for the expression (i.e., the transcription and translation) of a coding sequence in a host cell.

As used herein, a control sequence "directs the expression" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

As used herein, a "host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous DNA sequence.

As used herein, a cell has been "transformed" by exogenous DNA when such exogenous DNA has been introduced inside the cell membrane. Exogenous DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes and yeasts, for example, the exogenous DNA may be maintained on an episomal element, such as a plasmid. With respect to eukaryotic cells, a stably

transformed or transfected cell is one in which the exogenous DNA has become integrated into the chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the exogenous DNA.

As used herein, "transfection" or "transfected" refers to a process by which cells take up foreign DNA and integrate that foreign DNA into their chromosome. Transfection can be accomplished, for example, by various techniques in which cells take up DNA (e.g., calcium phosphate precipitation, electroporation, assimilation of liposomes, etc.) or by infection, in which viruses are used to transfer DNA into cells.

As used herein, a "target cell" is a cell that is selectively transfected over other cell types (or cell lines).

As used herein, a "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

As used herein, a "heterologous" region of a DNA construct is an identifiable segment of DNA within or attached to another DNA molecule that is not found in association with the other molecule in nature. Thus, when the heterologous region encodes a gene, the gene will usually be flanked by DNA that does not flank the gene in the genome of the source animal. Another example of a heterologous coding sequence is a construct where the coding sequence itself is not found in nature (e.g., synthetic sequences having codons different from the native gene). Allelic variation or naturally occurring mutational events do not give rise to a heterologous region of DNA, as used herein.

As used herein, a "modulator" of a polypeptide is a substance which can affect the polypeptide function.

An aspect of the present invention is isolated polynucleotides encoding a human Grb2-1 protein and substantially similar sequences exhibiting T-cell specificity. Isolated polynucleotide sequences are substantially similar if they are capable of hybridizing under moderately stringent conditions to SEQ ID NO:1 or they encode DNA sequences which are degenerate to SEQ ID NO:1 or are degenerate to those sequences capable of hybridizing under moderately stringent conditions to SEQ ID NO:1.

Moderately stringent conditions is a term understood by the skilled artisan and has been described in, for example, Sambrook *et al. Molecular Cloning: A Laboratory Manual*, 2nd edition, Vol. 1, pp. 101-104, Cold Spring Harbor Laboratory Press (1989). An exemplary hybridization protocol using moderately stringent conditions is as follows. Nitrocellulose filters are prehybridized at 65°C in a solution containing 6X SSPE, 5X Denhardt's solution (10g Ficoll, 10g BSA and 10g polyvinylpyrrolidone per liter solution), 0.05% SDS and 100 ug/ml tRNA. Hybridization probes are labeled, preferably radiolabelled (e.g., using the Bios TAG-IT® kit). Hybridization is then carried out for approximately 18 hours at 65°C. The filters are then washed twice in a solution of 2X SSC and 0.5% SDS at room temperature for 15 minutes. Subsequently, the filters are washed at 58°C, air-dried and exposed to X-ray film overnight at -70°C with an intensifying screen.

Degenerate DNA sequences encode the same amino acid sequence as SEQ ID NO:2 or the proteins encoded by that sequence capable of hybridizing under moderately stringent conditions to SEQ ID NO:1, but have variation(s) in the nucleotide coding sequences because of the degeneracy of the genetic code. For example, the

degenerate codons UUC and UUU both code for the amino acid phenylalanine, whereas the four codons GGX all code for glycine.

Alternatively, substantially similar sequences are defined as those sequences in which about 66%, preferably about 75% and most preferably about 90%, of the nucleotides or amino acids match over a defined length of the molecule. As used herein, substantially similar refers to the sequences having similar identity to the sequences of the instant invention. Thus nucleotide sequences that are substantially the same can be identified by hybridization or by sequence comparison. Protein sequences that are substantially the same can be identified by techniques such as proteolytic digestion, gel electrophoresis and/or microsequencing. Excluded from the definition of substantially similar sequences are the *grb* gene family members *grb2* and *grb3-3*.

Embodiments of the isolated polynucleotides of the invention include DNA, genomic DNA and RNA, preferably of human origin. A method for isolating a nucleic acid molecule encoding a Grb2-1 protein is to probe a genomic or cDNA library with a natural or artificially designed probe using art recognized procedures. See, e.g., "Current Protocols in Molecular Biology", Ausubel *et al.* (eds.) Greene Publishing Association and John Wiley Interscience, New York, 1989,1992. The ordinarily skilled artisan will appreciate that SEQ ID NO:1 or fragments thereof comprising at least 15 contiguous nucleotides are particularly useful probes. It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include, but are not limited to, radioisotopes, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes would enable the ordinarily skilled artisan are to isolate

complementary copies of genomic DNA, cDNA or RNA polynucleotides encoding Grb2-1 proteins from human, mammalian or other animal sources or to screen such sources for related sequences, e.g., additional members of the family, type and/or subtype, including transcriptional regulatory and control elements as well as other stability, processing, translation and tissue specificity-determining regions from 5' and/or 3' regions relative to the coding sequences disclosed herein, all without undue experimentation.

Another aspect of the invention is functional polypeptides encoded by the polynucleotides of the invention. An embodiment of a functional polypeptide of the invention is the human Grb2-1 protein having the amino acid sequence set forth in SEQ ID NO:2.

Another aspect of the invention is a method for preparing essentially pure human Grb2-1 protein. Yet another aspect is the human Grb2-1 protein produced by the preparation method of the invention. This protein has the amino acid sequence listed in SEQ ID NO:2 and includes variants with a substantially similar amino acid sequence that have the same function. The proteins of this invention are preferably made by recombinant genetic engineering techniques by culturing a recombinant host cell containing a vector encoding the polynucleotides of the invention under conditions promoting the expression of the protein and recovery thereof.

The isolated polynucleotides, particularly the DNAs, can be introduced into expression vectors by operatively linking the DNA to the necessary expression control regions, e.g., regulatory regions, required for gene expression. The vectors can be introduced into an appropriate host cell such as a prokaryotic, e.g., bacterial, or eukaryotic, e.g., yeast or mammalian cell by methods well known in the art. See Ausubel *et al.*, *supra*. The coding sequences for the desired proteins, having been prepared or isolated, can be cloned into any

suitable vector or replicon. Numerous cloning vectors are known to those of skill in the art and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include, but are not limited to, the bacteriophage (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), a baculovirus insect cell system, a *Drosophila* insect system, YCp19 (*Saccharomyces*) and pSV2neo (mammalian cells). See generally, "DNA Cloning": Vols. I & II, Glover *et al.* ed. IRL Press Oxford (1985) (1987); and T. Maniatis *et al.* ("Molecular Cloning" Cold Spring Harbor Laboratory (1982).

The gene can be placed under the control of control elements such as a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator, so that the DNA sequence encoding the desired protein is transcribed into RNA in the host cell transformed by a vector containing the expression construct. The coding sequence may or may not contain a signal peptide or leader sequence. The proteins of the present invention can be expressed using, for example, the *E. coli* tac promoter or the protein A gene (*spa*) promoter and signal sequence. Leader sequences can be removed by the bacterial host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437 and 4,338,397.

In addition to control sequences, it may be desirable to add regulatory sequences which allow for regulation of the expression of the protein sequences relative to the growth of the host cell. Regulatory sequences are known to those of skill in the art.

Exemplary are those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound or to various temperature or metabolic conditions. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

An expression vector is constructed so that the particular coding sequence is located in the vector with the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the control sequences being such that the coding sequence is transcribed under the "control" of the control sequences, i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding sequence. Modification of the sequences encoding the particular antigen of interest may be desirable to achieve this end. For example, in some cases it may be necessary to modify the sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame. The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector, such as the cloning vectors described above. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases, it may be desirable to produce mutants or analogues of human Grb2-1 protein. Mutants or analogues may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. See, e.g., T. Maniatis *et al.*, *supra*; "DNA Cloning,"

Vols. I and II, *supra*; and "Nucleic Acid Hybridization", *supra*.

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. Preferred mammalian cells include human embryonic kidney cells (293), monkey kidney cells, fibroblast (COS) cells, Chinese hamster ovary (CHO) cells, *Drosophila* or murine L-cells. If the expression system secretes the protein into growth media, the protein can be purified directly from the media. If the protein is not secreted, it is isolated from cell lysates or recovered from the cell membrane fraction. The selection of the appropriate growth conditions and recovery methods are within the skill of the art.

An alternative method to identify proteins of the present invention is by constructing gene libraries, using the resulting clones to transform *E. coli* and pooling and screening individual colonies using polyclonal serum or monoclonal antibodies to human Grb2-1.

The proteins of the present invention may also be produced by chemical synthesis such as solid phase peptide synthesis on an automated peptide synthesizer, using known amino acid sequences or amino acid sequences derived from the DNA sequence of the genes of interest. Such methods are known to those skilled in the art.

The proteins of the present invention or their fragments comprising at least one epitope can be used to produce antibodies, both polyclonal and monoclonal, directed to epitopes corresponding to amino acid sequences disclosed herein. If polyclonal antibodies are desired, a selected mammal such as a mouse, rabbit, goat or horse is immunized with a protein of the present invention, or its fragment, or a mutant protein. Serum

from the immunized animal is collected and treated according to known procedures. Serum polyclonal antibodies can be purified by immunoaffinity chromatography or other known procedures.

Monoclonal antibodies to the proteins of the present invention, and to the fragments thereof, can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by using hybridoma technology is well known. Immortal antibody-producing cell lines can be created by cell fusion and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA or transfection with Epstein-Barr virus. See, e.g., M. Schreier *et al.*, "Hybridoma Techniques" (1980); Hammerling *et al.*, "Monoclonal Antibodies and T-cell Hybridomas" (1981); Kennett *et al.*, "Monoclonal Antibodies" (1980); and U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,452,570; 4,466,917; 4,472,500; 4,491,632; and 4,493,890. Panels of monoclonal antibodies produced against the antigen of interest, or fragment thereof, can be screened for various properties, i.e., for isotype, epitope, affinity, etc. Monoclonal antibodies are useful in purification, using immunoaffinity techniques, of the individual antigens which they are directed against. Alternatively, genes encoding the monoclonals of interest may be isolated from the hybridomas by PCR techniques known in the art and cloned and expressed in the appropriate vectors. The antibodies of this invention, whether polyclonal or monoclonal have additional utility in that they may be employed as reagents in immunoassays, RIA, ELISA, and the like. The antibodies of the invention can be labeled with an analytically detectable reagent such as a radioisotope, fluorescent molecule or enzyme.

Chimeric antibodies, in which non-human variable regions are joined or fused to human constant regions

(see, e.g., Liu *et al.*, *Proc. Natl Acad. Sci. USA, 84,* 3439 (1987)), may also be used in assays or therapeutically. Preferably, a therapeutic monoclonal antibody would be "humanized" as described in Jones *et al.*, *Nature, 321,* 522 (1986); Verhoeyen *et al.*, *Science, 239,* 1534 (1988); Kabat *et al.*, *J. Immunol., 147,* 1709 (1991); Queen *et al.*, *Proc. Natl Acad. Sci. USA, 86,* 10029 (1989); Gorman *et al.*, *Proc. Natl Acad. Sci. USA, 88,* 34181 (1991); and Hodgson *et al.*, *Bio/Technology, 9:*, 421 (1991).

Another aspect of the present invention is modulators of the polypeptides of the invention. Functional modulation of Grb2-1 by a substance includes partial to complete inhibition of function, identical function, as well as enhancement of function. Embodiments of modulators of the invention include peptides, oligonucleotides and small organic molecules including peptidomimetics.

Another aspect of the invention is antisense oligonucleotides comprising a sequence which is capable of binding to the polynucleotides of the invention. Synthetic oligonucleotides or related antisense chemical structural analogs can be designed to recognize, specifically bind to and prevent transcription of a target nucleic acid encoding Grb2-1 protein by those of ordinary skill in the art. See generally, Cohen, J.S., *Trends in Pharm. Sci., 10,* 435(1989) and Weintraub, H.M., *Scientific American,* January (1990) at page 40.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates Grb2-1 protein function by affecting the binding of Grb2-1 protein to cellular binding partners. Examples of modulators include, but are not limited to peptides and small organic molecules including peptidomimetics. A Grb2-1 protein is provided having the amino acid sequence of human Grb2-1 (SEQ ID NO:2) or a functional derivative thereof together with a cellular

binding partner or synthetic analog thereof. The mixture is incubated with a test substance which is suspected of modulating Grb2-1 activity, under conditions which permit the formation of a Grb2-1 gene product/cellular binding partner complex. An assay is performed for the presence of the complex, free Grb2-1 protein or free cellular binding partner and the result compared to a control to determine the effect of the test substance.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates Grb2-1 activity by direct binding to Grb2-1 protein. Examples of modulators include, but are not limited to, peptides and small organic molecules including peptidomimetics. Modulator candidates are synthesized on a solid support by techniques such as those disclosed in Lam *et al.*, *Nature 354,* 82 (1991) or Burbaum *et al.*, *Proc. Natl. Acad. Sci. USA 92,* 6027 (1995) to provide solid support-associated modulator candidates. A labelled Grb2-1 protein is provided having the amino acid sequence of human Grb2-1 (SEQ ID NO:2) or a functional derivative thereof. Exemplary labels include directly attached fluorescent or colored dyes, biotin, radioisotopes or epitope tags, which are detectable by a suitable antibody. A mixture of solid support-associated modulator candidates and labelled Grb2-1 protein is incubated under conditions which can permit the formation of a Grb2-1 protein/modulator candidate complex. The solid support is separated from free soluble labelled Grb2-1 protein. An assay is performed for the presence of solid support-associated labelled protein. Solid supports complexed with labelled protein are isolated and the identity of the modulator candidate determined by techniques well known to those skilled in the art.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates Grb2-1 function by causing Grb2-1 to become

membrane bound resulting in the juxtaposition of Grb2-1 to downstream signalling proteins initiating signal transduction. A Grb2-1 protein is provided having the amino acid sequence of human Grb2-1 (SEQ ID NO:2) or a functional derivative thereof together with a cellular membrane. The mixture is incubated with a test substance which is suspected of modulating Grb2-1 activity, under conditions which permit the formation of a Grb2-1 gene product/cellular membrane complex. An assay is performed for the presence of the complex, free Grb2-1 protein or free cellular membrane and the result compared to a control to determine the effect of the test substance.

Modulation of Grb2-1 function would be expected to have effects on immune system function by affecting T-cell proliferation pathways. Any antagonist modulators so identified would be expected to have immunosuppressive activities and be useful as a therapeutic for the treatment and prevention of autoimmune diseases and transplant rejection. Any identifed agonist modulator would be expected to function as immunostimulators and be useful as a therapeutic for treatment of immune deficiency states such as HIV infection or cancer.

Further, Grb2-1 could be used to isolate proteins which interact with it and this interaction could be a target for interference. Inhibitors of protein-protein interactions between Grb2-1 and other factors could lead to the development of pharmaceutical agents for the modulation of Grb2-1 activity.

Methods to assay for protein-protein interactions, such as that of a Grb2-1 gene product/binding partner complex, and to isolate proteins interacting with Grb2-1 are known to those skilled in the art. Use of the methods discussed below enable one of ordinary skill in the art to accomplish these aims without undue experimentation.

The yeast two-hybrid system provides methods for detecting the interaction between a first test protein and a second test protein, *in vivo*, using reconstitution of the activity of a transcriptional activator. The method is disclosed in U.S. Patent No. 5,283,173; reagents are available from Clontech and Stratagene. Briefly, Grb2-1 cDNA is fused to a *Gal4* transcription factor DNA binding domain and expressed in yeast cells. cDNA library members obtained from cells of interest are fused to a transactivation domain of *Gal4*. cDNA clones which express proteins which can interact with Grb2-1 will lead to reconstitution of *Gal4* activity and transactivation of expression of a reporter gene such as *Gal1-lacZ*. Optionally, the host cells can be co-transfected with a protein tyrosine kinase to induce tyrosine phosphorylation of members of the cDNA library. Such phosphorylation is necessary for optimum interaction with the SH2 domain of grb2-1.

An alternative method is screening of λgt11, λZAP (Stratagene) or equivalent cDNA expression libraries with recombinant Grb2-1. Recombinant Grb2-1 protein or fragments thereof are fused to small peptide tags such as FLAG, HSV or GST. The peptide tags can possess convenient phosphorylation sites for a kinase such as heart muscle creatine kinase or they can be biotinylated. Recombinant Grb2-1 can be phosphorylated with ^{32}[P] or used unlabeled and detected with streptavidin or antibodies against the tags. λgt11cDNA expression libraries are made from cells of interest and are incubated with the recombinant Grb2-1, washed and cDNA clones isolated which interact with Grb2-1. See, e.g., T. Maniatis *et al, supra*.

Another method is the screening of a mammalian expression library in which the cDNAs are cloned into a vector between a mammalian promoter and polyadenylation site and transiently transfected in COS or 293 cells followed by detection of the binding protein 48 hours

later by incubation of fixed and washed cells with a labelled Grb2-1, prefereably iodinated, and detection of bound Grb2-1 by autoradiography (See Sims *et al.*, *Science 241*, 585-589 (1988) and McMahan *et al.*, *EMBO J. 10*, 2821-2832 (1991)). In this manner, pools of cDNAs containing the cDNA encoding the binding protein of interest can be selected and the cDNA of interest can be isolated by further subdivision of each pool followed by cycles of transient transfection, binding and autoradiography. Alternatively, the cDNA of interest can be isolated by transfecting the entire cDNA library into mammalian cells and panning the cells on a dish containing Grb2-1 bound to the plate. Cells which attach after washing are lysed and the plasmid DNA isolated, amplified in bacteria, and the cycle of transfection and panning repeated until a single cDNA clone is obtained (See Seed *et al*, *Proc. Natl. Acad. Sci. USA 84*, 3365 (1987) and Aruffo *et al.*, *EMBO J. 6*, 3313 (1987)). If the binding protein is secreted, its cDNA can be obtained by a similar pooling strategy once a binding or neutralizing assay has been established for assaying supernatants from transiently transfected cells. General methods for screening supernatants are disclosed in Wong *et al.*, *Science 228*, 810-815 (1985).

Another alternative method is isolation of proteins interacting with Grb2-1 directly from cells. Fusion proteins of Grb2-1 with GST or small peptide tags are made and immobilized on beads. Biosynthetically labeled or unlabeled protein extracts from the cells of interest are prepared, incubated with the beads and washed with buffer. Proteins interacting with Grb2-1 are eluted specifically from the beads and analyzed by SDS-PAGE. Binding partner primary amino acid sequence data are obtained by microsequencing. Optionally, the cells can be treated with agents that induce a functional response such as tyrosine phosphorylation of cellular proteins.

An example of such an agent would be a growth factor or cytokine such as interleukin-2.

Another alternative method is immunoaffinity purification. Recombinant Grb2-1 is incubated with labeled or unlabeled cell extracts and immunoprecipitated with anti-Grb2-1 antibodies. The immunoprecipitate is recovered with protein A-Sepharose and analyzed by SDS-PAGE. Unlabelled proteins are labeled by biotinylation and detected on SDS gels with streptavidin. Binding partner proteins are analyzed by microsequencing. Further, standard biochemical purification steps known to those skilled in the art may be used prior to microsequencing.

Yet another alternative method is screening of peptide libraries for binding partners. Recombinant tagged or labeled Grb2-1 is used to select peptides from a peptide or phosphopeptide library which interact with Grb2-1. Sequencing of the peptides leads to identification of consensus peptide sequences which might be found in interacting proteins.

Grb2-1 binding partners identified by any of these methods or other methods which would be known to those of ordinary skill in the art as well as those putative binding partners discussed above can be used in the assay method of the invention. Assaying for the presence of Grb2-1/binding partner complex are accomplished by, for example, the yeast two-hybrid system, ELISA or immunoassays using antibodies specific for the complex. In the presence of test substances which interrupt or inhibit formation of Grb2-1/binding partner interaction, a decreased amount of complex will be determined relative to a control lacking the test substance.

Assays for free Grb2-1 or binding partner are accomplished by, for example, ELISA or immunoassay using specific antibodies or by incubation of radiolabeled Grb2-1 with cells or cell membranes followed by

centrifugation or filter separation steps. In the presence of test substances which interrupt or inhibit formation of Grb2-1/binding partner interaction, an increased amount of free Grb2-1 or free binding partner will be determined relative to a control lacking the test substance.

Another aspect of the invention is pharmaceutical compositions comprising an effective amount of a Grb2-1 modulator of the invention and a pharmaceutically acceptable carrier. Pharmaceutical compositions of modulators of this invention for parenteral administration, i.e., subcutaneously, intramuscularly or intravenously or oral administration can be prepared.

The compositions for parenteral administration will commonly comprise a solution of the modulators of the invention or a cocktail thereof dissolved in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be employed, e.g., water, buffered water, 0.4% saline, 0.3% glycine and the like. These solutions are sterile and generally free of particulate matter. These solutions may be sterilized by conventional, well-known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, etc. The concentration of the modulator of the invention in such pharmaceutical formulation can vary widely, i.e., from less than about 0.5%, usually at or at least about 1% to as much as 15 or 20% by weight and will be selected primarily based on fluid volumes, viscosities, etc. according to the particular mode of administration selected.

Thus, a pharmaceutical composition of the modulator of the invention for intramuscular injection could be prepared to contain 1 mL sterile buffered water, and 50 mg of a protein of the invention. Similarly, a pharmaceutical composition of the modulator of the

invention for intravenous infusion could be made up to contain 250 ml of sterile Ringer's solution, and 150 mg of a modulator of the invention. Actual methods for preparing parenterally administrable compositions are well known or will be apparent to those skilled in the art and are described in more detail in, for example, *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania.

The physician will determine the dosage of the present therapeutic agents which will be most suitable and it will vary with the form of administration and the particular compound chosen, and furthermore, it will vary with the particular patient under treatment. Generally, the physician will wish to initiate treatment with small dosages substantially less than the optimum dose of the compound and increase the dosage by small increments until the optimum effect under the circumstances is reached. It will generally be found that when the composition is administered orally, larger quantities of the active agent will be required to produce the same effect as a smaller quantity given parenterally. The therapeutic dosage will generally be from 1 to 10 milligrams per day and higher although it may be administered in several different dosage units.

Depending on the patient condition, the pharmaceutical composition of the invention can be administered for prophylactic and/or therapeutic treatments. In therapeutic application, compositions are administered to a patient already suffering from a disease in an amount sufficient to cure or at least partially arrest the disease and its complications. In prophylactic applications, compositions containing the present compounds or a cocktail thereof are administered to a patient not already in a disease state to enhance the patient's resistance to the disease.

Single or multiple administrations of the pharmaceutical compositions can be carried out with dose

levels and pattern being selected by the treating physician. In any event, the pharmaceutical composition of the invention should provide a quantity of the modulators of the invention sufficient to effectively treat the patient.

Additionally, some diseases result from inherited defective genes. These genes can be detected by comparing the sequence of the defective gene with that of a normal one. Individuals carrying mutations in the Grb2-1 gene may be detected at the DNA level by a variety of techniques. Nucleic acids used for diagnosis (genomic DNA, mRNA, etc.) may be obtained from a patient's cells, such as from blood, urine, saliva or tissue biopsy, e.g., chorionic villi sampling or removal of amniotic fluid cells and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR, ligase chain reaction (LCR), strand displacement amplification (SDA), etc. prior to analysis. See, e.g., Saiki *et al.*, *Nature, 324*, 163-166 (1986), Bej, *et al.*, *Crit. Rev. Biochem. Molec. Biol., 26*, 301-334 (1991), Birkenmeyer *et al.*, *J. Virol. Meth., 35*, 117-126 (1991), Van Brunt, J., *Bio/Technology, 8*, 291-294 (1990)). RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid of the instant invention can be used to identify and analyze Grb2-1 mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal Grb2-1 genotype. Point mutations can be identified by hybridizing amplified DNA to rabiolabeled Grb2-1 RNA of the invention or alternatively, radiolabelled Grb2-1 antisense DNA sequences of the invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures (Tm). Such a diagnostic would be particularly useful for prenatal and even neonatal testing.

In addition, point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by yet other well-known techniques, e.g., direct DNA sequencing, single-strand conformational polymorphism. See Orita *et al.*, *Genomics*, *5*, 874-879 (1989). For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotides or by automatic sequencing procedures with fluorescent-tags. Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. The presence of nucleotide repeats may correlate to a causative change in Grb2-1 activity or serve as marker for various polymorphisms.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures. See, e.g., Myers *et al.*, *Science, 230,* 1242 (1985). In addition, sequence alterations, in particular small deletions, may be detected as changes in the migration pattern of DNA heteroduplexes in non-denaturing gel electrophoresis such as heteroduplex electrophoresis. See, e.g., Nagamine *et al.*, *Am. J. Hum. Genet., 45,* 337-339 (1989). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or

the chemical cleavage method as disclosed by Cotton *et al.* in *Proc. Natl. Acad. Sci. USA, 85,* 4397-4401 (1985).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization (e.g., heteroduplex electroporation, see, White *et al., Genomics, 12,* 301-306 (1992), RNAse protection (e.g., Myers *et al., Science, 230,* 1242 (1985)) chemical cleavage (e.g., Cotton *et al., Proc. Natl. Acad. Sci. USA, 85,* 4397-4401 (1985))), direct DNA sequencing, or the use of restriction enzymes (e.g., restriction fragment length polymorphisms (RFLP) in which variations in the number and size of restriction fragments can indicate insertions, deletions, presence of nucleotide repeats and any other mutation which creates or destroys an endonuclease restriction sequence). Southen blotting of genomic DNA may also be used to identify large (i.e., greater than 100 base pair) deletions and insertions.

In addition to conventional gel electrophoresis and DNA sequencing, mutations such as microdeletions, aneuploidies, translocations, inversions, can also be detected by *in situ* analysis. See, e.g., Keller *et al.*, DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA (1993). That is, DNA or RNA sequences in cells can be analyzed for mutations without isolation and/or immobilization onto a membrane. Fluorescence *in situ* hybridization (FISH) is presently the most commonly applied method and numerous reviews of FISH have appeared. See, e.g., Trachuck *et al., Science, 250,* 559-562 (1990), and Trask *et al., Trends, Genet., 7,* 149-154 (1991). Hence, by using nucleic acids based on the structure of the Grb2-1 genes, one can develop diagnostic tests for genetic mutations.

In addition, some diseases are a result of, or are characterized by, changes in gene expression which can be detected by changes in the mRNA. Alternatively, the Grb2-1 gene can be used as a reference to identify individuals expressing an increased or decreased level

of Grb2-1 protein, e.g., by Northern blotting or *in situ* hybridization.

Defining appropriate hybridization conditions is within the skill of the art. *See, e.g.*, "Current Protocols in Mol. Biol." Vol. I & II, Wiley Interscience. Ausbel *et al.* (eds.) (1992). Probing technology is well known in the art and it is appreciated that the size of the probes can vary widely but it is preferred that the probe be at least 15 nucleotides in length. It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioisotopes, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. As a general rule, the more stringent the hybridization conditions the more closely related genes will be that are recovered.

The putative role of Grb2-1 in signal transduction of the DNA synthesis pathway establishes yet another aspect of the invention which is gene therapy. "Gene therapy" means gene supplementation where an additional reference copy of a gene of interest is inserted into a patient's cells. As a result, the protein encoded by the reference gene corrects the defect and permits the cells to function normally, thus alleviating disease symptoms. The reference copy would be a wild-type form of the *grb2-1* gene or a gene encoding a protein or peptide which modulates the activity of the endogenous Grb2-1.

Gene therapy of the present invention can occur *in vivo* or *ex vivo*. *Ex vivo* gene therapy requires the isolation and purification of patient cells, the introduction of a therapeutic gene and introduction of the genetically altered cells back into the patient. A replication-deficient virus such as a modified retrovirus can be used to introduce the therapeutic

grb2-1 gene into such cells. For example, mouse Moloney leukemia virus (MMLV) is a well-known vector in clinical gene therapy trials. See, e.g., Boris-Lauerie *et al.*, *Curr. Opin. Genet. Dev.*, *3*, 102-109 (1993).

In contrast, *in vivo* gene therapy does not require isolation and purification of a patient's cells. The therapeutic gene is typically "packaged" for administration to a patient such as in liposomes or in a replication-deficient virus such as adenovirus as described by Berkner, K.L., in *Curr. Top. Microbiol. Immunol.*, *158*, 39-66 (1992) or adeno-associated virus (AAV) vectors as described by Muzyczka, N., in *Curr. Top. Microbiol. Immunol.*, *158*, 97-129 (1992) and U.S. Patent No. 5,252,479. Another approach is administration of "naked DNA" in which the therapeutic gene is directly injected into the bloodstream or muscle tissue. Another approach is administration of "naked DNA" in which the therapeutic gene is introduced into the target tissue by microparticle bombardment using gold particles coated with the DNA.

Cell types useful for gene therapy of the present invention include lymphocytes, hepatocytes, myoblasts, fibroblasts, any cell of the eye such as retinal cells, epithelial and endothelial cells. Preferably the cells are T lymphocytes drawn from the patient to be treated, hepatocytes, any cell of the eye or respiratory or pulmonary epithelial cells. Transfection of pulmonary epithelial cells can occur via inhalation of a neubulized preparation of DNA vectors in liposomes, DNA-protein complexes or replication-deficient adenoviruses. See, e.g., U.S. Patent No. 5,240,846.

Adaptor proteins such as grb2 function by juxtaposing cellular effectors with one another. For example, grb2 binds to motifs containing phosphotyrosine residues on activated growth factor receptors via its SH2 domain. The SH3 domains of grb2 bind the effector protein SOS, which, when in close proximity to membrane-

bound ras protein, transmits a signal by stimulating the formation of GTP-ras. Thus, the essential function of grb2 is to recruit SOS to the membrane so that it will encounter ras and activate it.

Activation of ras transmits a signal to turn on cell proliferation. In the case of T lymphocytes, such proliferation is essential for a cell-mediated immune response. Therefore, an agent that enhances proliferation of T-lymphocytes will have immunostimulant properties. Ideally, such an agent will specifically stimulate T-cells and not affect other cell types.

Thus, it will be appreciated that activation of the ras pathway in T-cells will be useful as a means of immunostimulation. Such activation may best be accomplished by causing a T-cell-specific adaptor protein such as grb2-1 to associate with the cell membrane, where it will bring SOS into close proximity with ras, causing activation. Association of grb2-1 with cell membranes may conveniently be accomplished by a substance capable of binding to the SH2 domain of grb2-1 and having membrane-seeking properties such as hydrophobic chemical moieties. Alternatively, the substance may be engineered with an acyl acceptor site such as a CAAX box or mimic thereof, which is capable of being acylated by cellular enzymes such as myristoyl or farnesyl transferases. Such a compound would be considered a prodrug.

Accordingly, another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates the ras pathway in T-lymphocytes by affecting the binding of Grb2-1 to the cell membrane. Membrane vesicles are prepared from cell suspensions by well known methods (for examples, see Vol. 31 of Methods in Enzymology). Any mammalian cells will be suitable, but T-lymphocytes are preferred. Exemplary is the Jurkat cell line obtainable from the American Type Culture Collection. The cells should be

deprived of growth factors for at least 18 hours prior to use. This may be accomplished by culturing the cells in serum-free medium. Optionally, the membranes may be treated with a phosphatase, e.g., potato acid phosphatase available from Sigma Chemicals to reduce background binding of Grb2-1. Grb2-1 protein is expressed as an epitope tagged fusion, e.g., GST-grb2-1, purified, e.g., by adsorption to glutathione-agarose and mixed with the membranes in a suitable buffer, e.g., 10 mM Tris pH 7.4, 150 mM NaCl. Test compounds are added and allowed to equilibrate for at least 1 hour at room temperature. Membrane vesicles are separated from the bulk solution, e.g., by filtration on 0.22 um membranes (Millipore) or by adsorption to wheat germ-conjugated agarose (Sigma Chemicals) or other methods known to those skilled in the art. The membranes are washed with buffer and assayed for bound Grb2-1 protein by addition of suitably labelled antibody to the epitope tag (anti-GST). The results are compared to control incubations lacking the compound to determine whether increased levels of Grb2-1 become associated with the membrane as a result of compound addition. For identification of prodrug compounds, activation is accomplished by adding either cell extracts or purified or recombinant acyl transferases to the membrane-grb2-1 incubation mixture, along with the appropriate substrates.

Another aspect of the invention is transgenic, non-human mammals capable of expressing the polynucleotides of the invention in any cell. Transgenic, non-human animals may be obtained by transfecting appropriate fertilized eggs or embryos of a host with the polynucleotides of the invention or with mutant forms found in human diseases. See, e.g., U.S. Patent Nos. 4,736,866; 5,175,385; 5,175,384 and 5,175,386. The resultant transgenic animal may be used as a model for the study of *grb2-1* gene function. Particularly useful transgenic animals are those which display a detectable

phenotype associated with the expression of the Grb2-1 protein. Drug development candidates may then be screened for their ability to reverse or exacerbate the relevant phenotype.

The present invention will now be described with reference to the following specific, non-limiting examples.

Example 1

Grb2-1 full-length cDNA Sequence Analysis

A search of a random cDNA sequence database consisting of short partial sequences known as expressed sequence tags (ESTs) with a murine *grb2* cDNA sequence reported by Suen, K. *et al.* in *Mol. Cell. Biol. 13,* 5500-5512 (1993) (SEQ ID NO: 3) disclosed an EST which matched the 5' end of the murine cDNA.
This cDNA was originally isolated from a human tonsils cDNA library.

Sequence analysis of the human *grb2-1* cDNA revealed a 700 nucleotide open reading frame (SEQ ID NO: 1) encoding a 217 amino acid protein (SEQ ID NO: 2) with a predicted molecular mass of 23.9 kDa, starting with an ATG at position 40 and terminating with a TGA at position 691 of SEQ ID NO: 1.

Alignment of the deduced amino acid sequence of human Grb2-1 (SEQ ID NO: 2) with the murine Grb2 amino acid sequence (SEQ ID NO: 4) was accomplished using the GCG program Bestfit. The overall amino acid identity was 60% with a two amino acid gap and is shown in Fig. 1 (top, human Grb2-1; bottom, murine Grb2).

Alignment of the deduced amino acid sequence of human Grb2-1 (SEQ ID NO: 2) with the human Grb2 amino acid sequence (SEQ ID NO: 5) was accomplished using the GCG program Bestfit. The overall amino acid identity was 58% with a two amino acid gap and is shown in Fig. 2 (top, human Grb2-1; bottom, human Grb2).

Example 2

Expression of Grb2-1

The complete coding sequence of Grb2-1 was amplified with the primers $^{5'}$gccccgagc*ggatcc*atggagtcc$^{3'}$ (SEQ ID NO: 6) and $^{5'}$ccgccg*ctcgag*tcacaggtgcac$^{3'}$ (SEQ ID NO: 7) to introduce *Bam*HI and a *Xho*I endonuclease restriction sites. These were used to clone the restricted DNA fragment into the corresponding sites of the plasmid pGEX-4T-1 (Pharmacia). The vector encompassing the GST-Grb2-1 fusion protein was transformed into *E.coli* XL-1 blue (Stratagene). Protein expression was stimulated with IPTG (0.5 mM final concentration) for 4 hrs. Cells were resuspended in 50 mM Tris/HCl (pH 8.0), 100 mM NaCl, 1 mM EDTA, 0.4 mg/ml lysozyme and 0. 5 mM PMSF and lysed with 2 mg/ml deoxycholic acid and 1.4 % sarcosyl. DNA was sheared by sonication. After pelleting the particulate fraction, Triton X-100 was added to a final 1.8% concentration. The fusion protein was affinity purified over glutathione-agarose beads (Molecular Probes).

Example 3

Tissue Distribution of Grb2-1

Northern blots of tissue mRNA were conducted to determine the tissue distribution of *grb2-1* gene transcription. The endonucleases *Not*I and *Xho*I were used to excise a 2 kb insert from the pBluescript vector. (See Example 2). The liberated cDNA fragment was isolated using low-melting temperature 1% agarose gel electrophoresis and radiolabelled with [^{32}P]-dATP using a randomly primed labelling kit from Stratagene.

Membranes containing mRNA from multiple human tissues (Clontech) were hybridized with the probes and washed under high stringency conditions. Hybridized mRNA was visualized using a phosphorimager (Molecular Dynamics). The results in Fig. 3 show that the 3.5 kb

Grb2 transcript is expressed at roughly equivalent levels in all tissues examined, whereas the expression of the 2.1 kb Grb2-1 transcript is restricted to lymphoid tissues. The following two-letter abbreviations are used:

He	HEART
Br	BRAIN
Pl	PLACENTA
Lu	LUNG
Li	LIVER
Mu	MUSCLE
Ki	KIDNEY
Pa	PANCREAS
Sp	SPLEEN
Th	THYMUS
Pr	PROSTATE
Te	TESTIS
Ov	OVARY
In	INTESTINE
Co	COLON
Pb	PBL

The present invention may be embodied in other specific forms without departing from the spirit or essential attributes thereof, and, accordingly, reference should be made to the appended claims, rather than to the foregoing specification, as indicating the scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Dunnington, Damien D
Ni, Jian
Shoelson, Steven E

(ii) TITLE OF INVENTION: Growth Factor Receptor-Binding Protein 2 Homolog

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SmithKline Beecham Corporation
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19406

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE: 04-DEC-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Baumeister, Kirk
(B) REGISTRATION NUMBER: 33,833
(C) REFERENCE/DOCKET NUMBER: P50416P

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 610-270-5096
(B) TELEFAX: 610-270-5090

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 700 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGAGCTGA GCCCAGCTGC TGGAGCCCCG AGCAGCGGCA TGGAGTCCGT GGCCCTGTAC 60

AGCTTTCAGG CTACAGAGAG CGACGAGCTG GCCTTCAACA AGGGAGACAC ACTCAAGATC 120

CTGAACATGG AGGATGACCA GAACTGGTAC AAGGCCGAGC TCCGGGGTGT CGAGGGATTT 180

ATTCCCAAGA ACTACATCCG CGTCAAGCCC CATCCGTGGT ACTCGGGCAG GATTTCCCGG 240

CAGCTGGCCG AAGAGATTCT GATGAAGCGG AACCATCTGG GAGCCTTCCT GATCCGGGAG 300

AGTGAGAGCT CCCCAGGGGA GTTCTCTGTG TCTGTGAACT ATGGAGACCA GGTGCAGCAC 360

TTCAAGGTGC TGCGTGAGGC CTCGGGGAAG TACTTCCTGT GGGAGGAGAA GTTCAACTCC 420

CTCAACGAGC TGGTCGACTT CTACCGCACC ACCACCATCG CCAAGAAGCG GCAGATCTTC 480

CTGCGCGACG AGGAGCCCTT GCTCAAGTCA CCTGGGGCCT GCTTTGCCCA GGCCCAGTTT 540

GACTTCTCAG CCCAGGACCC CTCGCAGCTC AGCTTCCGCC GTGGCGACAT CATTGAGGTC 600

CTGGAGCGCC CAGACCCCCA CTGGTGGCGG GGCCGGTCCT GCGGGCGCGT TGGCTTCTTC 660

CCACGGAGTT ACGTGCAGCC CGTGCACCTG TGAGCAGCCC 700

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 217 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Glu Ser Val Ala Leu Tyr Ser Phe Gln Ala Thr Glu Ser Asp Glu
1 5 10 15

Leu Ala Phe Asn Lys Gly Asp Thr Leu Lys Ile Leu Asn Met Glu Asp
20 25 30

Asp Gln Asn Trp Tyr Lys Ala Glu Leu Arg Gly Val Glu Gly Phe Ile
35 40 45

Pro Lys Asn Tyr Ile Arg Val Lys Pro His Pro Trp Tyr Ser Gly Arg
50 55 60

Ile Ser Arg Gln Leu Ala Glu Glu Ile Leu Met Lys Arg Asn His Leu
65 70 75 80

Gly Ala Phe Leu Ile Arg Glu Ser Glu Ser Ser Pro Gly Glu Phe Ser
85 90 95

Val Ser Val Asn Tyr Gly Asp Gln Val Gln His Phe Lys Val Leu Arg
100 105 110

Glu Ala Ser Gly Lys Tyr Phe Leu Trp Glu Glu Lys Phe Asn Ser Leu
115 120 125

Asn Glu Leu Val Asp Phe Tyr Arg Thr Thr Thr Ile Ala Lys Lys Arg
130 135 140

Gln Ile Phe Leu Arg Asp Glu Glu Pro Leu Leu Lys Ser Pro Gly Ala
145 150 155 160

Cys Phe Ala Gln Ala Gln Phe Asp Phe Ser Ala Gln Asp Pro Ser Gln
165 170 175

Leu Ser Phe Arg Arg Gly Asp Ile Ile Glu Val Leu Glu Arg Pro Asp
180 185 190

Pro His Trp Trp Arg Gly Arg Ser Cys Gly Arg Val Gly Phe Phe Pro
195 200 205

Arg Ser Tyr Val Gln Pro Val His Leu
210 215

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2362 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CGGCTGAGCC CAGAGCTCTT CAGCAGGGAA GATTCCCTTC CCCCCTGCTT CAGGCTGCTG 60

AGCACTGAGC GGCGCTCAGA ATGGAAGCCA TCGCCAAATA TGACTTCAAA GCTACTGCTG 120

ACGATGAGCT GAGCTTCAAA AGGGGGGACA TCCTTAAGGT TTTGAATGAA GAATGTGACC 180

AGAACTGGTA TAAGGCAGAA CTCAATGGGA AAGATGGCTT CATCCCCAAG AATTACATAG 240

AAATGAAACC ACATCCGTGG TTTTTTGGCA AAATCCCCAG AGCCAAGGCA GAAGAAATGC 300

TCAGCAAACA GCGGCATGAC GGGGCCTTCC TGATCCGAGA GAGCGAGAGC GCTCCTGGGG 360

ACTTCTCCCT GTCCGTCAAG TTTGGAAATG ATGTGCAGCA CTTCAAGGTG CTCCGCGACG 420

GAGCCGGGAA GTATTTCCTG TGGGTGGTGA AGTTTAATTC TTTGAATGAG CTGGTAGATT 480

ACCACAGATC AACATCCGTG TCCAGGAACC AGCAGATATT CTTACGGGAC ATAGAACAGA 540

TGCCACAGCA GCCAACCTAC GTCCAGGCGC TCTTTGACTT TGACCCCCAG GAGGATGGCG 600

AGCTGGGCTT TCGCAGAGGA GACTTCATTC ATGTCATGGA TAACTCAGAT CCCAATTGGT 660

GGAAAGGGGC CTGCCACGGG CAGACCGGCA TGTTTCCCCG CAATTATGTC ACCCCAGTGA 720

ACCGGAACGT CTAAGAAGCA AAAGAGATTA TTTAAAGAAA GTGAAAAGTT AAGACCGTTC 780

ACAAGAATTA CACCCACACG CTGCCTGTCA CAGCCTGTGA GGGAACGCAG AACACCTGGC 840

TGGGTCCCAC GGGTGACCCT CTCATTGGGT TGCAACTTTG GGGGGTGGGG TAGGGGTGTT 900

TGATTTCATA ATGCCAAAAC TTAACCTATT GAATGAATTA CAGTTTTTAT TACGGAATCT 960

CGCCGCTACC CCTGTTCCCC TCCTGTGTCC TTTTTCTCGT TCTTTCTTTC CTGTCCAGTG 1020

CATGATGTTT AAGGCCACAT ATAGTCCAGC TGATGCCAAT AATAAAAGAC AAGAAACCAA 1080

GTGGGCTGGT ATTTTCTCTA TGCAAAATGT CTGTGGAGAT GGATGGACTG AAAGAGCCGG 1140

ATTCCTCACA CACAGGGGCA GCCAGTGCTT CTGGGGCCCT GGTTGGGGTT CACCCGAGAT 1200

GCCCAGGGGT ACCGCCTCCA GCCTCAGGCC TGGAGCATTC CATCAAAGTT GGAATTAGGG 1260

GAAGGAGGCC CACTGACCTC CCCGGTCTCC TGAGAGTCAG ACTGCAGGCC CTCCCCTCTC 1320

CCACTGCTTC CCTTCAGGTG TTTTGACGTT TTTTTGTTTG TTTGTTTGTT TTTTTTAAAT 1380

AGTGCCTTTG TCTTATTTCA AGGGTGTTCA TAAATGGTAT TTGTACCATT TTTTTTTTTA 1440

ATAAGTTAAA GACAGTCCAG AGCTTTTCAG TTGATTCGTC TCCTATCCTG TGTAAATATT 1500

TTCCTCTCAG GGCAGGGGAA AGAGGACAGA GAAAGGAGCT GGTAGAAGCA GAGAGTGTAT 1560

TTCCCATCTT GAATGGGCCG GAGGTCTCGA AGCCTCAGCT TTTACTTTGT GAGCTGCAAC 1620

ACTCGCTTCG ACTCAGACTC AGTGGACATC AGAGTCTCTG TCTCCGTGTC TCAGTTATGG 1680

TCTGCTCTCT CTATGCCTGG AGCTACTGAT CCAAACACAA GACGGTCAGA GGAGCCCTGG 1740

CATCCTTCAC TCTTATAGGC CTACATGCGA TGGGCTTGAA GAGAATTGGC CTTCATTTTT 1800

CACGCTCGAT TTCCCCACCT GTCCAAGGGT GCAGATAGCA GGATTGTGGG GTCGGCCGCA 1860

TCTCTCTTGG CAGTGGGTTG TGCTGCCCCT CCCCCCACCC CACCCCTCCA TCCCACAAAG 1920

CCTTTCGAGG TAGGAGAGAC CAAGACCAGT GCACAGCTTT TAACTCTTGT CTTCTGTGAG 1980

TTTTCCGTCC TCTGTAGTCA CGTGCCTGCA CACCGTTCTC TCCACCCTGC CTCCTTCCCA 2040

CAGAGAAGCA GGGCTCCACC CAGGCCTTCC CTTGGAGTCG GTGCATCCAT GGGCTGCTAG 2100

ACTCTTGCGG GTAGAGTCTC CCCTGGACTT AGCATTGTGA GATGGACTCA GCTCAGGGCG 2160

CCCCTAAGGC TCGGAGGCCA TGGCCTTCTG GGTCCCCACC TCTCCTGGTT CGCTGCTGCT 2220

CTCCCTGCTG ATGATAAAGT AATCTCTGGA GTCACACCTG GGCCATGTGA TTGTTTTTAT 2280

TTTGGAATTG GAGATATCAT GAAGCCTTGC TGAACTAAGT TTTGTGTGTA TATATTTAAA 2340

GATCAGTGTT TAAATAAAAA AA 2362

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 217 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Glu Ala Ile Ala Lys Tyr Asp Phe Lys Ala Thr Ala Asp Asp Glu
1 5 10 15

Leu Ser Phe Lys Arg Gly Asp Ile Leu Lys Val Leu Asn Glu Glu Cys
20 25 30

Asp Gln Asn Trp Tyr Lys Ala Glu Leu Asn Gly Lys Asp Gly Phe Ile
35 40 45

Pro Lys Asn Tyr Ile Glu Met Lys Pro His Pro Trp Phe Phe Gly Lys
50 55 60

Ile Pro Arg Ala Lys Ala Glu Glu Met Leu Ser Lys Gln Arg His Asp
65 70 75 80

Gly Ala Phe Leu Ile Arg Glu Ser Glu Ser Ala Pro Gly Asp Phe Ser
85 90 95

Leu Ser Val Lys Phe Gly Asn Asp Val Gln His Phe Lys Val Leu Arg
100 105 110

Asp Gly Ala Gly Lys Tyr Phe Leu Trp Val Val Lys Phe Asn Ser Leu
115 120 125

Asn Glu Leu Val Asp Tyr His Arg Ser Thr Ser Val Ser Arg Asn Gln
130 135 140

Gln Ile Phe Leu Arg Asp Ile Glu Gln Met Pro Gln Gln Pro Thr Tyr
145 150 155 160

Val Gln Ala Leu Phe Asp Phe Asp Pro Gln Glu Asp Gly Glu Leu Gly
165 170 175

Phe Arg Arg Gly Asp Phe Ile His Val Met Asp Asn Ser Asp Pro Asn
180 185 190

Trp Trp Lys Gly Ala Cys His Gly Gln Thr Gly Met Phe Pro Arg Asn
195 200 205

Tyr Val Thr Pro Val Asn Arg Asn Val
210 215

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 217 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Glu Ala Ile Ala Lys Tyr Asp Phe Lys Ala Thr Ala Asp Asp Glu
1 5 10 15

Leu Ser Phe Lys Arg Gly Asp Ile Leu Lys Val Leu Asn Glu Glu Cys
20 25 30

Asp Gln Asn Trp Tyr Lys Ala Glu Leu Asn Gly Lys Asp Gly Phe Ile
35 40 45

Pro Lys Asn Tyr Ile Glu Met Lys Pro His Pro Trp Phe Phe Gly Lys
50 55 60

Ile Pro Arg Ala Lys Ala Glu Glu Met Leu Ser Lys Gln Arg His Asp
65 70 75 80

Gly Ala Phe Leu Ile Arg Glu Ser Glu Ser Ala Pro Gly Asp Phe Ser
85 90 95

Leu Ser Val Lys Phe Gly Asn Asp Val Gln His Phe Lys Val Leu Arg
100 105 110

Asp Gly Ala Gly Lys Tyr Phe Leu Trp Val Val Lys Phe Asn Ser Leu
115 120 125

Asn Glu Leu Val Asp Tyr His Arg Ser Thr Ser Val Ser Arg Asn Gln
130 135 140

Gln Ile Phe Leu Arg Asp Ile Glu Gln Val Pro Gln Gln Pro Thr Tyr
145 150 155 160

Val Gln Ala Leu Phe Asp Phe Asp Pro Gln Glu Asp Gly Glu Leu Gly
165 170 175

Phe Arg Arg Gly Asp Phe Ile His Val Met Asp Asn Ser Asp Pro Asn
180 185 190

Trp Trp Lys Gly Ala Cys His Gly Gln Thr Gly Met Phe Pro Arg Asn
195 200 205

Tyr Val Thr Pro Val Asn Arg Asn Val
210 215

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

GCCCCGAGCG GATCCATGGA GTCC 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

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

CCGCCGCTCG AGTCACAGGT GCAC 24

CLAIMS

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding human Grb2-1 having the nucleotide sequence as set forth in SEQ ID NO:1 from nucleotide 40 to 690;

(b) a polynucleotide capable of hybridizing to the complement of a polynucleotide according to (a) under moderately stringent hybridization conditions and which encodes a functional human Grb2-1; and

(c) a degenerate polynucleotide according to (a) or (b).

2. An isolated polynucleotide having the nucleotide sequence as set forth in SEQ ID NO:1.

3. A functional polypeptide encoded by the polynucleotide of claim 1.

4. The functional polypeptide of claim 3 which is human Grb2-1 having the amino acid sequence set forth in SEQ ID NO:2.

5. The polynucleotide of claim 1 which is DNA.

6. The polynucleotide of claim 5 which is genomic DNA.

7. The polynucleotide of claim 1 which is RNA.

8. A vector comprising the DNA of claim 5.

9. A recombinant host cell comprising the vector of claim 8.

10. A method for preparing essentially pure human Grb2-1 protein comprising culturing the recombinant host cell of claim 9 under conditions promoting expression of the protein and recovering the expressed protein.

11. Human Grb2-1 produced by the process of claim 10.

12. An antisense oligonucleotide comprising a sequence which is capable of binding to the polynucleotide of claim 1.

13. A modulator of the polypeptide of claim 3.

14. The modulator of claim 13 which is a peptide.

15. The modulator of claim 13 which is a small organic molecule.

16. The small organic molecule of claim 15 which is a peptidomimetic.

17. A method for assaying a medium for the presence of a substance that modulates Grb2-1 activity by affecting the binding of Grb2-1 to cellular binding partners comprising the steps of:

(a) providing a Grb2-1 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular binding partner or synthetic analog thereof;

(b) incubating with a test substance which is suspected of modulating Grb2-1 activity under conditions which permit the formation of a Grb2-1 protein/cellular binding partner complex;

(c) assaying for the presence of the complex, free Grb2-1 protein or free cellular binding partner; and

(d) comparing to a control to determine the effect of the substance.

18. Grb2-1 protein modulating compounds identified by the method of claim 17.

19. A method for assaying a medium for the presence of a substance that modulates Grb2-1 activity by direct binding to Grb2-1 protein comprising the steps of:

(a) providing a labelled Grb2-1 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof

(b) providing solid support-associated modulator candidates;

(c) incubating a mixture of the labelled Grb2-1 protein with the support-associated modulator candidates under conditions which can permit the formation of a Grb2-1 protein/modulator candidate complex;

(d) separating the solid support from free soluble labelled Grb2-1 protein;

(e) assaying for the presence of solid support-associated labelled protein;

(f) isolating the solid support complexed with labelled Grb2-1 protein; and

(g) identifying the modulator candidate.

20. A method for assaying a medium for the presence of a substance that modulates Grb2-1 activity by causing Grb2-1 to become membrane bound comprising the steps of:

(a) providing a Grb2-1 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular membrane;

(b) incubating with a test substance which is suspected of modulating Grb2-1 activity under conditions which permit the formation of a Grb2-1 protein/cellular membrane complex;

(c) assaying for the presence of the membrane-bound complex, free Grb2-1 protein or free cellular membrane; and

(d) comparing to a control to determine the effect of the substance.

21. Grb2-1 protein modulating compounds identified by the method of claim 20.

22. A method for the treatment of a patient having need to modulate Grb2-1 activity comprising administering to the patient a therapeutically effective amount of the modulating compound of claims 18 or 20.

23. A pharmaceutical composition comprising the modulating compound of claims 18 or 21 and a pharmaceutically acceptable carrier.

24. A method of diagnosing conditions associated with Grb2-1 protein deficiency which comprises:

(a) isolating a polynucleotide sample from an individual;

(b) assaying the polynucleotide sample and a polynucleotide encoding Grb2-1 having the nucleotide sequence as set forth in SEQ ID NO:1 from nucleotide 40 to 690; and

(c) comparing differences between the polynucleotide sample and the Grb2-1 polynucleotide, wherein any differences indicate mutations in the Grb2-1 gene.

25. A method of treating conditions which are related to insufficient Grb2-1 protein function which comprises:

(a) isolating cells from a patient deficient in Grb2-1 protein function;

(b) altering the cells by transfecting the polynucleotide of claim 1 into the cells wherein a Grb2-1 protein is expressed; and

(c) introducing the cells back to the patient to alleviate the condition.

26. A method of treating conditions which are related to insufficient Grb2-1 protein function which comprises administering the polynucleotide of claim 1 to a patient deficient in Grb2-1 protein function wherein a Grb2-1 protein is expressed and alleviates the condition.

27. A method for assaying a medium for the presence of a substance that modulates the ras pathway in T-lymphocytes by affecting the binding of Grb2-1 to the cell membrane comprising:

(a) providing a Grb2-1 protein having the amino acid sequence of Grb2-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular membrane vesicle;

(b) incubating with a test substance which is suspected of affecting the binding of Grb2-1 to the cell membrane under conditions which permit the formation of a Grb2-1 protein/cellular membrane vesicle complex;

(c) assaying for the presence of the membrane-bound complex, free Grb2-1 protein or free cellular membrane; and

(d) comparing to a control to determine the effect of the substance.

28. Ras pathway modulating compounds identified by the method of claim 27.

29. A transgenic non-human animal capable of expressing in any cell thereof the DNA of claim 5.

```
          .         .         .         .         .
  1 MESVALYSFQATESDELAFNKGDTLKILNMEDDQNWYKAELRGVEGFIPK 50
    ||.:| |.|.||..|||.|.:||.||:|| | |||||||||.| :|||||
  1 MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPK 50
          .         .         .         .         .
 51 NYIRVKPHPWYSGRISRQLAEEILMKRNHLGAFLIRESESSPGEFSVSVN 100
    ||| :|||||: |:|.|. |||:| |..| ||||||||||.||:||:||.
 51 NYIEMKPHPWFFGKIPRAKAEEMLSKQRHDGAFLIRESESAPGDFSLSVK 100
          .         .         .         .         .
101 YGDQVQHFKVLREASGKYFLWEEKFNSLNELVDFYRTTTIAKKRQIFLRD 150
    :|::|||||||::.|||||| ||||||||||:.|.|.:.:..||||||
101 FGNDVQHFKVLRDGAGKYFLWVVKFNSLNELVDYHRSTSVSRNQQIFLRD 150
          .         .         .         .         .
151 EEPLLKSPGACFAQAQFDFSAQDPSQLSFRRGDIIEVLERPDPHWWRGRS 200
     |.: . |.  :.|| |||.:|:.::|:|||||:|.|::..||:||:| :
151 IEQMPQQPT..YVQALFDFDPQEDGELGFRRGDFIHVMDNSDPNWWKGAC 198
          .
201 CGRVGFFPRSYVQPVHL.. 217
     |..|:|||.|| ||:
199 HGQTGMFPRNYVTPVNRNV 217
```

FIGURE 1

```
                .         .         .         .         .
  1 MESVALYSFQATESDELAFNKGDTLKILNMEDDQNWYKAELRGVEGFIPK 50
    ||.:| |.|.||..|||.|.:||.||:|| | |||||||||.| :|||||
  1 MEAIAKYDFKATADDELSFKRGDILKVLNEECDQNWYKAELNGKDGFIPK 50
            .         .         .         .         .
 51 NYIRVKPHPWYSGRISRQLAEEILMKRNHLGAFLIRESESSPGEFSVSVN 100
       ||| :|||||: |:|.|. |||:| |..| ||||||||||.||:||:||.
 51 NYIEMKPHPWFFGKIPRAKAEEMLSKQRHDGAFLIRESESAPGDFSLSVK 100
                .         .         .         .         .
101 YGDQVQHFKVLREASGKYFLWEEKFNSLNELVDFYRTTTIAKKRQIFLRD 150
       :|::|||||||||::.||||||  ||||||||||:.|.|.:.:..||||||
101 FGNDVQHFKVLRDGAGKYFLWVVKFNSLNELVDYHRSTSVSRNQQIFLRD 150
                .         .         .         .         .
151 EEPLLKSPGACFAQAQFDFSAQDPSQLSFRRGDIIEVLERPDPHWWRGRS 200
       |.: . |.  :.|| |||.:|:.::|:|||||:|.|:..||:||:| :
151 IEQVPQQPT..YVQALFDFDPQEDGELGFRRGDFIHVMDNSDPNWWKGAC 198
                .
201 CGRVGFFPRSYVQPVHL* 217
       |..|:|||.|| ||:
199 HGQTGMFPRNYVTPVNRNV* 217
```

FIGURE 2

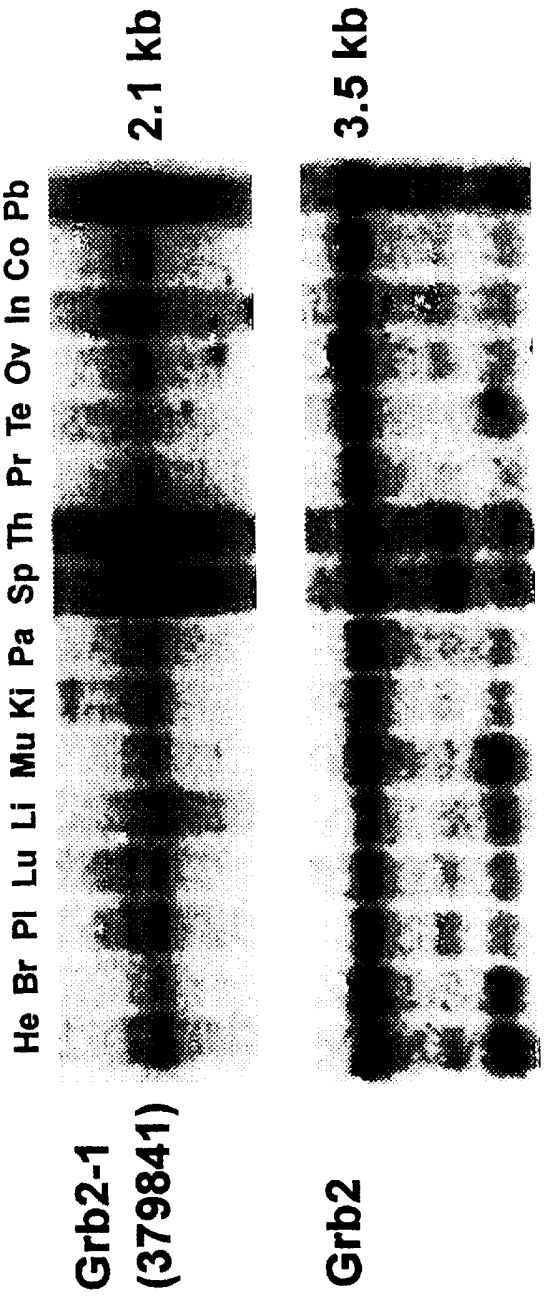
2.1 kb
3.5 kb
He Br Pl Lu Li Mu Ki Pa Sp Th Pr Te Ov In Co Pb
Grb2-1 (379841)
Grb2

Figure 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15883

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 38/00, 48/00; C07H 21/00; C07K 14/00; C12N 15/00; C12P 21/00

US CL :Please See Extra Sheet.

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)

U.S. : 424/93.21; 435/69.1, 320.1; 514/2, 44; 530/350; 536/23.1, 24.5; 800/2

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

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

APS, MEDLINE, EMBASE, BIOSIS, CAPLUS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,434,064 (SCHLESSINGER et al.) 18 July 1995, see entire document.	1-28
Y	MOLECULAR BIOLOGY OF THE CELL, Volume 4, issued November 1993, Stern et al, "The Human GRB2 and Drosophilia Drk Genes Can Functionally Replace the Caenorhabditis elegans Cell Signaling Gene sem-5", pages 1175-1188, see entire document.	1-28
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, Volume 89, issued October 1992, Matuoka et al, "Cloning of ASH, a ubiquitous protein composed of one Src homology region (SH) 2 and two SH3 domains, from human and rat cDNA libraries", pages 9015-9019, see entire document.	1-28

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	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
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"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
"E"	earlier document published on or after the international filing date	"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
"L"	document which may throw doubts 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		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 APRIL 1996	22 APR 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer D. CURTIS HOGUE, JR. Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15883

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SCIENCE, Volume 264, issued 13 MAY 1994, Fath et al, "Cloning of a Grb2 Isoform with Apoptotic Properties", pages 971-974, see entire document.	1-28
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, Volume 89, issued October 1992, Margolis et al, "High-efficiency expression/cloning of epidermal growth factor-receptor-binding proteins with Src homology 2 domains", pages 8894-8898, see entire document.	1-28
Y	MOLECULAR AND CELLULAR BIOLOGY, Volume 13, No. 9, issued September 1993, Suen et al, "Molecular Cloning of the Mouse grb2 Gene: Differential Interaction of the Grb2 Adaptor Protein with Epidermal Growth Factor and Nerve Growth Factor Receptors", pages 5500-5512, see entire document.	1-28
Y	FEBS LETTERS, Volume 369, issued 1995, Chardin et al, "The Grb2 adaptor", pages 47-51, see entire document.	1-28
Y	CURRENT BIOLOGY, Volume 3, issued 15 May 1995, Cowburn, "Src homology adaptor proteins: more than the sum of the parts", pages 429-430, see entire document.	1-28
Y	GENOMICS, Volume 22, issued 1994, Yulug et al, "Mapping GRB2, a Signal Transduction Gene in the Human and the Mouse", pages 313-318, see entire document.	1-28
Y	FEBS LETTERS, Volume 338, issued 1994, Downward, "The GRB2/Sem-5 adaptor protein", pages 113-117, see entire document.	1-28
Y	GENE, Volume 134, issued 1993, Wasenius et al, "Sequence of a chicken cDNA encoding a GRB2 protein", pages 299-300, see entire document.	1-28
Y	CELL, Volume 70, issued 07 August 1992, Lowenstein et al, "The SH2 and SH3 Domain-Containing Protein GRB2 Links Receptor Tyrosine Kinases to ras Signaling", pages 431-442, see entire document.	1-28

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15883

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

424/93.21; 435/69.1, 320.1; 514/2, 44; 530/350; 536/23.1, 24.5; 800/2