PCT WORLD INTELLECTUAL PRO International

INTERNATIONAL APPLICATION PUBLISHED UNDE WO 9606168A2

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47, G01N 33/68, C07K 16/18, C12Q 1/68 — A2

(11) International Publication Number: **WO 96/06168**

(43) International Publication Date: 29 February 1996 (29.02.96)

(21) International Application Number: PCT/US95/10574

(22) International Filing Date: 18 August 1995 (18.08.95)

(30) Priority Data:

08/292,683	18 August 1994 (18.08.94)	US
08/399,411	6 March 1995 (06.03.95)	US

(71) Applicant: LA JOLLA CANCER RESEARCH FOUNDATION [US/US]; 10901 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventor: HUANG, Shi; 3623 Caminito Carmel Landing, San Diego, CA 92130 (US).

(74) Agents: WILSON, Barry, S. et al.; Campbell and Flores, Suite 700, 4370 La Jolla Village Drive, San Diego, CA 92122 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, 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), ARIPO patent (KE, MW, SD, SZ, UG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: RETINOBLASTOMA PROTEIN-INTERACTING ZINC FINGER PROTEINS

(57) Abstract

The present invention provides mammalian retinoblastoma (Rb)protein-interacting zinc finger (RIZ) proteins, nucleic acid molecules encoding the RIZ and antibodies specific for the RIZ. The invention also provides screening assays for identifying an agent that effectively alters the association of a RIZ with a second molecule, which can bind to the RIZ. The invention also provides active fragments of RIZ containing the PR domain, which can regulate transcription. In addition, the invention provides methods for introducing a nucleic acid molecule encoding a RIZ into a cell and for contacting a cell with an effective agent in order to modulate a function of a cell. Such methods are useful, for example, for inducing growth of a cardiac cell or a neuronal cell in a subject and for effecting normal growth control to a tumor cell or causing differentiation of tumor cells. The invention further provides methods for detecting a RIZ in a sample by detecting the RIZ or a nucleic acid molecule encoding the RIZ. Such methods are useful to diagnose a pathology that is characterized by an increased or decreased level of a RIZ in a cell or by expression of a mutant RIZ in a cell.

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.

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

RETINOBLASTOMA PROTEIN-INTERACTING ZINC FINGER PROTEINS

This invention was made in part with government support under CA57496, awarded by the National Institutes of Health, and 5T30 CA30199, awarded by the Cancer Center Core. The United States government has certain rights in this invention.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention relates to the field of molecular biology and, in particular, to nucleic acid molecules encoding an Rb-interacting zinc finger (RIZ) protein and a conserved domain of a RIZ protein that is involved in regulating gene transcription.

BACKGROUND INFORMATION

The retinoblastoma Rb protein is known to play a key role in controlling normal cell proliferation and differentiation. The ability of a cell to divide requires the cell to pass through the various phases of

the cell cycle. Although Rb is believed to keep normal cells from dividing by maintaining them in a phase of the cell cycle known as G_1 or G_0, the precise mechanism underlying Rb function is unknown. It is known, however, that Rb can bind various cellular proteins, includingproteins involved in regulating gene transcription. Thus, Rb may exert its action by interacting with such cellular proteins.

The role that Rb plays in controlling cell growth makes it an attractive target for promoting the growth of tissues that normally do not grow because of the action of Rb. For example, cardiac muscle tissue and nervous tissue that have lost function due to cell death are not usually repaired by subsequent proliferation of the remaining live cells. Thus, a method to block the growth controlling function of Rb can be useful for inducing tissue repair in situations of cardiac or neural cell death.

Rb also is known as a tumor suppressor since the abnormal growth of a cancer cell can result from inactivation of Rb protein. Such inactivation can occur either due to a mutation or to inactivation of Rb protein subsequent to binding a viral oncoprotein, a product of an oncogenic tumor virus. A particular region in Rb called the Rb pocket appears to be critical for its growth controlling function since Rb inactivation by mutation or by oncoprotein binding impacts this region.

The importance of the Rb pocket in the functioning of Rb and the understanding that viral oncoproteins can regulate Rb by binding the pocket suggest that there may be normal cellular proteins that can regulate the function of Rb by binding the pocket. The identification of such proteins will provide new approaches to regulate the control of cell proliferation mediated by Rb in diseases such as those that involve loss of cardiac or neural function or in the control of cancer.

Thus, a need exists to identify proteins that can bind to and regulate Rb in order to provide new approaches for controlling cell proliferation and differentiation. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

The present invention provides substantially purified mammalian Rb-interacting zinc finger proteins (RIZ), including for example, human RIZ and rat RIZ. In addition, the invention provides active fragments of a RIZ such as the sequences EIRCEEKPEDL (SEQ ID NO: 6) and EIRCDEKPEDL (SEQ ID NO: 91), which bind Rb. The invention also provides antibodies that can specifically bind to a RIZ or a mutant RIZ.

The invention further provides nucleic acid molecules encoding mammalian RIZ and active fragments thereof, vectors containing the nucleic acid molecules

and host cells containing the vectors. In addition, the invention provides nucleotide sequences that can specifically hybridize to a nucleic acid molecule encoding a RIZ or a mutant nucleic acid molecule encoding a RIZ.

In addition, the invention provides a peptide comprising a PR domain, which is conserved among various proteins and can be involved in regulating the transcription of a target gene. In general, a PR domain peptide contains about 100 to about 120 amino acids that characteristically are arranged as a series of three highly conserved sequences of about ten to about twelve amino acids each, which are separated from each other by less conserved sequences of about 24 to about 34 amino acids each. A PR domain of the invention is exemplified by the PR domain present in the RIZ protein disclosed herein and by the PR domain present in the proteins PRDI-BF1, Evi-1 and egl-43.

The present invention further provides fusion proteins comprising a PR domain of the invention operably linked to a peptide that can bind to a DNA regulatory element. For example, a fusion protein of the invention can comprise a PR domain operably linked to a peptide that binds to a particular gene promotor or enhancer, wherein binding of the fusion protein to a target gene, which is a gene containing the particular promotor or enhancer, can alter expression of the target gene. Thus, a fusion protein of the invention can be useful for regulating the transcription of one or more target genes.

In addition, the invention further provides methods of identifying transcription factors and oncogenic proteins that bind a PR domain peptide or a RIZ active fragment containing a PR domain. The identification of such factors and proteins provides new approaches to manipulate cell differentiation and transformation.

The invention also provides a screening assay useful for identifying agents that can effectively alter the association of a RIZ with a second molecule such as Rb or can effectively alter the activity of a RIZ. By altering the association of a RIZ with a second molecule or altering the activity of a RIZ, an effective agent can modulate a function of a cell such as cell proliferation.

The invention further provides methods for promoting the growth of a cell such as a neural cell or cardiac muscle cell by contacting the cell with an effective agent. For example, cell growth can be promoted by introducing into a cell an effective agent such as an expression vector having an expression control sequence operably linked to a nucleotide sequence encoding an active fragment of a RIZ, wherein the active fragment lacks the growth-suppressing properties of a complete RIZ protein. In addition, the invention provides methods for restoring normal controlled cell growth to cancer cells by introducing into the cancer cells an expressible nucleic acid molecule encoding a complete RIZ protein.

The invention also provides methods of detecting a RIZ in a sample by detecting the presence of the RIZ protein or of a nucleic acid molecule encoding the RIZ. Such methods can be used to diagnose a pathology characterized by an increased or decreased level of expression of a RIZ in a cell or by expression of a mutant RIZ. Such a method also can be used to diagnose a pathology characterized by a mutant nucleic acid molecule encoding a RIZ.

The invention further provides methods useful for isolating Rb tumor suppressor protein or a mutant Rb from a sample. For example, Rb can be isolated from a sample by affinity chromatography using a RIZ or a RIZ active fragment such as the sequences EIRCEEKPEDL (SEQ ID NO: 6) or EIRCDEKPEDL (SEQ ID NO: 91).

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A and 1B show alternative nucleotide sequences (SEQ ID NOS: 1 and 94) and the deduced amino acid (a.a.) sequence (SEQ ID NO: 2) of full-length rat RIZ protein.

Figure 1A shows the cDNA sequence (SEQ ID NO: 1) and the deduced a.a. sequence of full-length rat RIZ protein (SEQ ID NO: 2). Numbers at right indicate nucleotide position; numbers at left indicate amino acid position. The following features are underlined: an upstream in-frame stop codon (nucleotide position 100-102), a cr2 core motif (a.a. positions 304-309), 8 zinc

fingers (a.a. positions 357-377, 478-499, 387-407, 1125-1203 (finger 4-6), 1323-1343 and 1445-1466) and a putative nuclear localization signal (a.a. positions 867-874). A putative leucine zipper is located from a.a. position 667-695; T, H and L residues within the zipper are underlined. Single letter amino acid symbols are used.

Figure 1B shows an alternative nucleotide sequence (SEQ ID NO: 94), which is present at the 5'-end of a nucleotide sequence encoding rat RIZ (SEQ ID NO: 2). The alternative nucleotide sequence (SEQ ID NO: 94) replaces nucleotides 1 to 91 at the 5'-end of the nucleotide sequence shown in Figure 1A (SEQ ID NO: 1).

Figures 2A to 2C show homologies between rat RIZ and various other proteins. Single letter amino acid symbols are used. Numbers indicate amino acid positions in relation to the complete protein.

Figure 2A compares RIZ amino acid sequences with various E1A sequences. E1A sequences of the different strains of adenoviruses are from Kimelman et al., J. Virol. 53:399-409 (1985), Moran and Mathews, Cell 48:177-178 (1987), and Ishino et al., Virology 165:95-102 (1988). Identical or closely related residues are boxed. Single letter amino acid symbols are used. Sequence domains, RIZ cr1 (SEQ ID NO: 79), Ad2E1A cr1 (SEQ ID NO: 44), Ad5 cr1 (SEQ ID NO: 45), Ad7 cr1 (SEQ ID NO: 46), Ad12 cr1 (SEQ ID NO: 47), EA7 cr1 (SEQ ID NO: 48), Ad40 cr1 (SEQ ID NO: 49), RIZ cr2 (SEQ ID NO: 65), Ad2E1A cr2

(SEQ ID NO: 66), Ad5 cr2 (SEQ ID NO: 67), Ad7 cr2 (SEQ ID NO: 68), Ad12 cr2 (SEQ ID NO: 69), EA7 cr2 (SEQ ID NO: 70), Ad40 cr2 (SEQ ID NO: 71), RIZ ce1 (SEQ ID NO: 72), Ad2E1A ce1 (SEQ ID NO: 73), Ad5 ce1 (SEQ ID NO: 74), Ad7 ce1 (SEQ ID NO: 75), Ad12 ce1 (SEQ ID NO: 76), EA7 ce1 (SEQ ID NO: 77) and Ad40 ce1 (SEQ ID NO: 78) are shown.

Figure 2B shows RIZ putative SH3 and SH3-binding domains. Panel a: Sequence comparison of RIZ with other known SH3 domain-containing proteins (Lowenstein et al., Cell 70:431-442 (1992)). Identical or closely related residues are boxed and the phosphate-binding loop in RIZ (SEQ ID NO: 80) is underlined. Sequences from GRB2 N-terminus (SEQ ID N0: 50), GRB2 C-terminus (SEQ ID NO: 51), P85 (SEQ ID NO: 52), v-abl (SEQ ID NO: 53), c-src (SEQ ID NO: 54), GAP (SEQ ID NO: 55), PLC (SEQ ID NO: 56) and v-crk (SEQ ID NO: 57) are shown. Panel b: A RIZ putative SH3-binding motif compared with SH3 motifs from known SH3-binding proteins (Ren et al., Science 259: 1157-1161 (1993)). Identical or closely related residues are boxed. Sequences from (SEQ ID NO: 81) Formin (SEQ ID NO: 58), 3BP1 (SEQ ID NO: 59), 3BP2 (SEQ ID NO: 60) and m4mAChR (SEQ ID NO: 61) are shown.

Figure 2C shows homology between RIZ and PRDI-BF1 proteins (Keller and Maniatis, Genes Devel. 5: 868-879 (1991)). Panel a: Alignment of RIZ zinc fingers 4 to 6 (SEQ ID NO: 82) with PRDI-BF1 zinc fingers 1 to 3 (SEQ ID NO: 62). A consensus sequence is shown with nonidentical residues indicated by dots. Potential DNA

contact residues are marked by stars (Pavletich and Pabo, Science 252:809-817 (1991)). Panel b: Amino terminal homology between RIZ (a.a. position 39-115: SEQ ID NO: 83 and a.a. position 116-145: SEQ ID NO: 84) and PRDI-BF1 (a.a. position 60-139: SEQ ID NO: 63 and a.a. position 140-169: SEQ ID NO: 64)). A consensus sequence is shown with nonidentical residues indicated by dots.

Figure 3: Schematic representation of RIZ domain structure. PR: domain homologous to PRDI-BF1; AR: acidic region or E1A-related region; LZ: leucine-zipper; cr1 and cr2: conserved regions 1 and 2; ce1: common epitope 1. Zinc (Zn-) fingers, GTPase and SH3 and SH3-binding domains also are shown.

Figure 4 demonstrates that ^{35}S-labeled RIZ (a.a. position 1-575) specifically binds to Rb *in vitro*. Binding assays were performed in the absence or presence of peptides derived from RIZ or SV40 large T antigen. T-pep: peptide of large T antigen (a.a. position 101-118); T-pep*: single amino acid residue mutant of T-pep (107^{Glu}); RIZ-pep: peptide of RIZ (a.a. position 304-314); RIZ-pep*: single amino acid mutant of RIZ pep (307^{Gly}).

Figure 5A and 5B: Use of Rb deletion mutants to map the RIZ binding site of Rb.

Figure 5A: Purified glutathionine S-transferase fusion protein containing a C-terminal RIZ fragment (a.a.

position 245-573) was tested for binding to ^{35}S-labeled Rb wild-type (wt-Rb) and to various deletion mutants (lanes 2-5) as shown in Figure 5B. Wild-type (wt) full length Rb (A9), Cys to Phe mutation of full length Rb (H209), Rb deletion mutants from amino acid positions 515-619 (NM), 585-697 (PP) and 804-928 (B3) are shown.

Figure 5B: Schematic map of Rb wild-type (wt-Rb) and Rb deletion mutants. The two sub-domains of the Rb pocket are represented by black boxes. Mutants that bind a glutathionine S-transferase (GST) RIZ (a.a. position 245-573) are indicated by a "+" sign. p56 Rb: 56 kD fragment of Rb from a.a. position 379-928.

Figure 6 shows *in vitro* binding of Rb from HT1080 cells with purified rat RIZ (a.a. position 215-462) fused C-terminal to glutathionine S-transferase (GST). The positions of hypophosphorylated Rb (pRb) and phosphorylated Rb (ppRb) from HT1080 cell extract are indicated in lane 4. Numbers to left indicate the migration of molecular weight markers (kiloDaltons).

Figures 7A to 7C demonstrate that rat RIZ protein binds DNA.

Figure 7A: SDS-PAGE (10% acrylamide) and Coomassie blue staining of GST; GSTZ13: GST-RIZ (a.a. position 245-573 containing zinc fingers 1-3); and GSTZ46: GST-RIZ (a.a. position 1114-1260 containing zinc fingers 4 to 6). KD indicates the migration of molecular weight markers.

Figure 7B: Binding of ^{32}P-labeled rat genomic DNA to GST, GSTZ13 and GSTZ46 in the presence of zinc ions.

Figure 7C: As in Figure 7B, except zinc ions were not added.

Figures 8A and 8B show the GTP-binding activity of rat RIZ GTPase domain (a.a. position 760-949).

Figure 8A: SDS-PAGE (10% acrylamide) and Coomassie blue staining of purified GST-G: GST-RIZ (760-949: RIZ GTPase domain fused C-terminal to glutathionine S-transferase).

Figure 8B: ^{32}P-GTP binding by GST (lane 1) and GST-G (lanes 2-6). Binding conducted in the absence or presence (lanes 3-6) of excess unlabeled nucleotides as indicated.

Figures 9A to 9C show alternative nucleotide sequences (SEQ ID NOS: 3 and 95-96), the deduced amino acid sequence (SEQ ID NO: 4) of full-length human RIZ and, additional 3' untranslated sequence of human RIZ cDNA (SEQ ID NO: 97).

Figure 9A shows a complete nucleotide sequence (SEQ ID NO: 3) and the deduced amino acid sequence (SEQ ID NO: 4) of full-length human RIZ. Three letter amino acid symbols are used. Numbers at right indicate the nucleotide position.

Figure 9B shows two alternative nucleotide sequences (SEQ ID NOS: 95-96), which replace the first 129 nucleotides present at the 5'-end of the nucleotide sequence shown in Figure 9A (SEQ ID NO: 3). The encoded amino acid sequences for clones 5Y and 1Y are shown (SEQ ID NOS: 102 and 103, respectively).

Figure 9C presents 3' untranslated sequence of human RIZ cDNA including the poly A tail (SEQ ID NO: 97). The sequence is located in human RIZ cDNA at the far 3' end of the cDNA, downstream of the 3' untranslated sequence shown in Figure 9A (SEQ ID NO: 3). Additional untranslated sequence is present between the 3' end of SEQ ID NO: 3 and the 5' end of SEQ ID NO: 97.

Figure 10 compares the complete human RIZ amino acid sequence (indicated as hRIZ; SEQ ID NO: 4) with the complete rat RIZ amino acid sequence (indicated as rRIZ; SEQ ID NO: 2). A consensus sequence is shown. Single letter amino acid symbols are used. Amino acids that are identical in hRIZ and rRIZ are shown as a ".".

Figures 11A and 11B show RIZ mRNA expression in adult and fetal rat tissues, as indicated. Relative amounts of RNA loaded were compared by probing for Actin (see bottom of each blot). Numbers to the left of each figure indicate position of molecular weight markers as indicated (Kb: kilobases).

Figure 11A presents a northern blot of adult mRNA probed with ^{32}P-labeled rat RIZ (1.9 Kb fragment

representing a.a. position 245-883). Att-20 is a mouse pituitary cell line.

Figure 11B presents an RNase protection experiment using RNA from a 16 day fetal rat (E16) and from an adult rat probed with ^{32}P-labeled rat RIZ (representing a.a. position 463-574).

Figure 12 presents an RNase protection experiment using mRNA from various human cell lines and from a rat cell line (indicated as GC). The 321 base pair marker (321 b) indicates protection of ^{32}P-labeled rat RIZ (a.a. position 457-579) while the 130 base pair marker (130 b) indicates protection of Actin.

Figure 13 presents an alignment of the PR domain peptides present in four transcription regulatory proteins. The sequences are aligned so as to maximize homology. A "-" indicates a space introduced to maintain homology. Numbers across the top of the sequences indicates the amino acid number in the PR domain peptide. Numbers at the beginning of each sequence indicate the amino acid position from the N-terminus of each respective protein. Identical or similar residues found at a particular position in at least three of the four peptides are shaded. Identical amino acids at a position in the PR domain peptide are indicated by black shading and conserved regions are indicated by gray shading. "PRDI-BF1" indicates human positive regulatory domain 1-binding factor 1 (SEQ ID NO: 98); "RIZ" indicates human RIZ protein (SEQ ID NO: 99); "EVI1" indicates human

ecotropic viral integration site-1 myeloid transforming gene product (SEQ ID NO: 100); "egl-43" indicates the *C. elegans* egl-43 gene product (SEQ ID NO: 101).

Figure 14 shows the nucleotide sequence (SEQ ID NO: 104) and the deduced a.a. sequence (SEQ ID NO: 105) of coding exons 4-6 of mouse RIZ. Numbers at right indicate nucleotide position.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a novel mammalian Rb-interacting zinc finger protein, designated RIZ. RIZ is a nuclear phosphoprotein that acts as a cell differentiation factor. RIZ can modulate a function of a cell by binding to retinoblastoma (Rb) protein, which is involved in regulating cell proliferation. In addition, RIZ can act to regulate transcription.

Rb is a nuclear phosphoprotein of 110 kiloDaltons (kD) that can bind DNA and is expressed in all tissue types examined thus far. The complete absence of Rb function is associated with the development of childhood retinoblastoma. In addition, Rb is mutated in a variety of cancer types, including various carcinomas and sarcomas, indicating a role for Rb in oncogenesis. Expression of exogenous Rb in various types of tumor cells suppresses the tumor phenotype (for review, see Lee et. al., J. Cell Biochem. 38:213-227 (1988)).

The function of Rb at the biochemical level in a cell is poorly understood. Rb is present in phosphorylated and unphosphorylated forms in the cell. The phosphorylation status of Rb oscillates during the cell cycle with the hypophosphorylated form correlating with the maintenance of the cell in G_1 phase of the cell cycle. Thus, the state of phosphorylation plays an important role in Rb function.

Rb protein binds to several DNA tumor viral oncoproteins, including the adenoviral E1A protein, the SV40 large T antigen and the E7 protein of the human papilloma virus (DeCaprio et al., Cell 54:275-283 (1988); Whyte et al., Cell 56:67-75 (1989); Dyson et al., Science 243:934-937 (1989)). The oncoproteins E1A and large T antigen bind to a similar region of Rb protein known as the Rb pocket, which is formed by two non-contiguous amino acid sequences in the protein (Hu et al., EMBO J. 9:1147-1155 (1990); Huang et al., EMBO J. 9:1815-1822 (1990); Kaelin et al., Mol. Cell. Biol. 10:3761-3769 (1990), each of which is incorporated herein by reference). The binding to Rb by these viral oncoproteins can alter normal Rb function.

Rb also can bind various cellular factors, including, for example, c-Myc and N-Myc (Rustgi et al., Nature 352:541-544 (1991)), E2F (Bagchi et al., Cell 65:1063-1072 (1991)), activating transcription factor-2 (ATF-2; Kim et al., Nature 358:331-334 (1992)), c-Abl (Welch and Wang, Cell 75:779-790 (1993), MyoD (Gu et al.,

Cell 72:309-324 (1993)) and brahma-related gene-1 (BRG-1; Dunaief et al. Cell 79:119-130 (1994)). Since these cellular factors are involved, for example, in gene regulation and cell differentiation, Rb can have a role in regulating the activity of cell transcription and differentiation factors.

As disclosed herein, RIZ is a normal cellular protein that binds to the Rb pocket. RIZ binding to Rb is unlike that of an oncoprotein since RIZ functions as a differentiation factor that helps to maintain cells in the G_0 or G_1 phases of the cell cycle. This is based on the fact that RIZ can bind to Rb in the cell, the latter being a known regulator of cell proliferation and differentiation, and that RIZ is structurally related to a known differentiation and transcription factor PRD1-BF1/Blimp-1 (Huang, Cell 78:9 (1994)).

The ability to regulate cell growth has important implications for various human diseases or conditions. Cancer is an example of a disease that results from a breakdown in the ability of a cell to regulate its growth. In contrast, there are examples such as cardiac muscle cells and neural cells where the maintenance of cell growth control contributes to a sustained loss in organ or tissue function following a disease or injury that resulted in cell death. In these situations, the compromised tissue or organ fails to regenerate fully because the remaining live cells are incapable of undergoing proliferation to replace the lost function.

Heart disease provides an example where cardiac muscle cell death due to ischemia or other injury results in a loss of heart function. Generally, proliferation of the remaining live cardiac cells to regenerate the lost cardiac muscle function does not occur in adults. Although myocardial cell proliferation can occur during embryonic and neonatal development, this capacity to proliferate is lost soon after birth. In a similar manner, neural damage resulting from trauma or disease is not usually followed by regeneration of neural function because the remaining neural cells are maintained in the G_1 phase of the cell cycle. Transcriptional regulators such as Rb play an important role in controlling whether cells can enter the cell cycle and proliferate. In contrast, inactivation of Rb is involved in the unregulated growth of a cancer cell.

As disclosed herein, RIZ can bind to Rb and can regulate the ability of Rb to maintain cells in the G_1 phase of the cell cycle. Methods that affect the ability of Rb and RIZ to associate or that affect the activity of a RIZ can be used to modulate cell proliferation. RIZ can regulate the growth of normal adult cardiac muscle cells by preventing the cells from proliferating following cardiac muscle cell death. RIZ can function to maintain cells in G_1 by interacting with Rb through the cr2 domain of RIZ. In addition, the functional differentiation state of a cell, which involves maintenance of a cell in G_1, is affected, in part, through the action of other RIZ domains such as the PR domain, GTP binding domain and the zinc finger domains.

The present invention provides a substantially purified RIZ protein. In a particular embodiment, the invention provides substantially purified mammalian RIZ. The invention provides, for example, human RIZ having substantially the amino acid sequence shown in Figure 9A (SEQ ID NO: 4) or Figure 9A with the first three amino acids replaced by SEQ ID NO: 102 or 103 (shown in Figure 9B), rat RIZ having substantially the amino acid sequence shown in Figure 1A (SEQ ID NO: 2) and mouse RIZ having substantially the amino acid sequence shown in Figure 14 (SEQ ID NO: 105).

As used herein, the term "substantially the amino acid sequence" means a sequence that is similar to the disclosed amino acid sequence. For example, an amino acid sequence that is substantially similar to human RIZ (SEQ ID NO: 4) or to rat RIZ (SEQ ID NO: 2) can have one or more amino acid additions, deletions or substitutions that do not substantially alter the ability of the encoded protein to function like a RIZ. In view of this definition, it should be recognized, for example, that the rat RIZ sequence shown in Figure 1A (SEQ ID NO: 2), which is 84% homologous to the human RIZ sequence has substantially the amino acid sequence of human RIZ (SEQ ID NO: 4). Similarly, the rat RIZ cr2 fragment sequence EIRCEEKPEDL (SEQ ID NO: 6) is substantially the sequence of the human RIZ cr2 fragment sequence, EIRCDEKPEDL (SEQ IN NO: 91). The latter two sequences differ by a single conservative substitution of a Glu in the rat for an Asp in the human in the residue following the Cys.

As used herein, the term "substantially purified" means a protein that is in a form that is relatively free from contaminating lipids, proteins, nucleic acids or other cellular material normally associated with a protein in a cell. A substantially purified human RIZ protein can be obtained, for example, using well known biochemical methods of purification or by expressing a recombinant nucleic acid molecule encoding a RIZ such as the nucleic acid molecule shown as SEQ ID NO: 3. In addition, an amino acid sequence consisting of at least a portion of the amino acid sequence of SEQ ID NO: 4 can be chemically synthesized or can be produced by expressing a portion of the nucleotide sequence shown as SEQ ID NO: 3.

As used herein, the terms "protein" or "polypeptide" are used in the broadest sense to mean a sequence of amino acids that can be encoded by a cellular gene or by a recombinant nucleic acid sequence or can be chemically synthesized. In some cases, the term "polypeptide" is used in referring to a portion of an amino acid sequence encoding a full length protein. An active fragment of a RIZ is an example of such a polypeptide. A protein can be a complete, full length gene product, which can be a core protein having no amino acid modifications, or can be a post-translationally modified form of a protein such as a phosphoprotein, glycoprotein, proteoglycan, lipoprotein or nucleoprotein.

The full length rat RIZ protein contains 1706 amino acids and has a calculated molecular mass of 187,437 Daltons (Figure 1A; SEQ ID NO: 2). The rat RIZ contains a 6 residue E1A related motif (a.a. position 304-309) known as the cr2 core motif, which is related to the LXCXE (SEQ ID NO: 5) core motif of E1A. Additional E1A related motifs in RIZ include the cr1 motif and a C-terminal motif designated "conserved epitope 1" (ce1) because of its antigenic relationship to a homologous motif in the C-terminus of E1A (see Example II). Rat RIZ also contains 8 zinc fingers, a putative GTPase domain, a putative leucine zipper and a putative nuclear localization signal (Figures 1 and 3).

All three E1A-related motifs in rat RIZ are located in an acidic region that consists of about 150 residues (AR; Figure 3) and resembles a highly acidic region in the E1A 12S protein (Moran and Matthews, Cell 48:177-178 (1987)). In both RIZ and E1A, the related motifs are arranged in the same order and the spacing between cr1 and cr2 is similar. However, the ce1 motif is located much closer to cr2 in RIZ than in E1A (see Figure 2A).

The rat RIZ protein sequence contains known GTPase motifs (Table 1) organized in an orderly fashion and separated by consensus spacings (Bourne et al., Nature 349:117-127 (1991)). The G1 or Walker type-A motif ($GX_4GKX_7(I/V)$; SEQ ID NO: 14), which represents the phosphate-binding loop (P-loop), occurs at a.a. position

749 in RIZ and identifies a guanine or adenine nucleotide-binding site (Walker et al., EMBO J. 1: 945-951 (1982); Saraste et al., Trends Biochem. Sci. 15: 430-434. (1990)). The sequence around residue 749 also is similar to the src homology 3 (SH3) domain conserved in many non-receptor tyrosine kinases and other proteins (Figure 2B panel a); Pawson and Gish, Cell 71:359-362 (1992)). RIZ also contains a proline-rich region that has several potential SH3-binding motifs (Figure 2B, panel b); Renet al., Science, 259:1157-1161 (1993)).

Table 1 Putative GTPase Domain in RIZ

	G1	G2	G3	G4
Consensus	GXXXXGK$^{S}_{T}$	D(X)$_N$T	DXXG	$^{NK}_{TC}$XD
	*(22)		(23)	(24)
RIZ	^{749}GKPNDGKA	^{785}DERET	^{853}DSEG	^{912}TQPD
	(85)	(86)	(89)	(90)
		^{796}D(x)$_{12}$T		
		(87)		
		^{821}D(X)$_{11}$T		
		(88)		
FtsZ	^{106}GGTGTGAA	^{122}DLGILT	^{180}DAFG	^{295}TSLD
	(25)	(26)	(29)	(31)
		^{158}DSLIT	^{253}DLSG	
		(27)	(30)	
		^{212}DVRT		
		(28)		
CDC42	^{10}GDGAVGKT	^{32}YVPT	^{57}DTAG	^{115}TQID
	(32)	(33)	(34)	(35)
DOG-SR2	^{419}GVNGVGKS	^{455}DT	^{516}DTAG	^{584}TKFD
	(36)		(34)	(37)
EF-Tu	^{13}GHVDHGKT	^{50}D(X)$_{10}$T	^{80}DCPG	^{135}NKCD
	(38)	(39)	(40)	(41)
Ha-Ras	^{10}GAGGVGKS	^{33}DPT	^{57}DTAG	^{116}NKCD
	(42)	(43)	(34)	(41)

Comparison of the putative G1-G4 GTPase domains in the RIZ protein sequence with the conserved sequence motifs in the GTPase superfamily (single letter code and X is any residue, Bourne et al., 1991) For reference to the listed sequences (except RIZ and FtsZ) see Bourne et al. (1991) For reference to FtsZ, see RayChaudhuri and Park, Nature 359:251-254,(1992).

* Number in parenthesis below each sequence indicates SEQ ID NO:.

Sequence homology shows that a mammalian RIZ protein contains eight zinc-finger motifs organized as two widely separated clusters in the N-terminal (fingers 1 to 3) and C-terminal (fingers 4 to 6) regions (Figure 3). A search of the National Biomedical Research Foundation protein database revealed that the most significant homology for zinc fingers was for RIZ fingers 4 to 6, which are about 39 % (33 out of 85) identical to fingers 1 to 3 of the human transcriptional repressor PRDI-BF1 (see Figure 2C; Keller and Maniatis, *supra*, 1991).

RIZ also contains a region of about 100 residues near the N-terminus that is designated "PR" because it is 42% homologous with a similar N-terminal region from PRDI-BF1 (see Figure 2C) and Blimp-1 (Huang, *supra*, 1994). PR, also referred to herein as the "PR domain peptide", is homologous to an N-terminal portion of the mammalian Evi-1 protein (Morishita et al., Cell 54:831-840 (1988); Morishita et al. Oncogene 5:936-971 (1990)) and to an N-terminal portion of the *C. elegans* egl-43 protein, which is a homolog of Evi-1 (Garriga et al., Genes Devel. 7:2097-2109 (1993); see Figure 13).

The Evi-1 protein is the product of the ecotropic viral integration site-1 myeloid transforming gene. Aberrant expression of the Evi-1 gene occurs in

human acute myelodysplastic leukemia (AML), myelodysplastic diseases (MDS) and chronic myelocytic leukemia (CML), due to translocations or inversions involving chromosome band 3q26 (see, for example, Morishita et al., Proc. Natl. Acad. Sci., *USA* 89:3937-3941 (1992)). The first 102 amino acids of the Evi-1 protein shown in Figure 13 (SEQ ID NO: 100), are encoded by nucleotides previously reported to be 5'-untranslated sequence (Morishita et al., *supra*, 1990); the methionine present in position designated 108 in Figure 13 (i.e., the 103rd amino acid shown in Figure 13 for Evi-1) was assigned as the start codon. However, as disclosed herein, designation of an ATG codon further upstream as the start codon elucidates blocks A and B of the PR domain peptide of Evi-1 as shown in Figure 13 (SEQ ID NO: 100).

In general, a PR domain peptide is about 100 to about 120 amino acids in length and contains three highly conserved sequences, designated blocks A, B and C, which consist of about 10 to about 12 amino acids, separated by less conserved sequences of about 20 to about 35 amino acids (see Figure 13). Each of blocks A, B and C for RIZ, Evi-1 and egl-43 are encoded by an individual exon. The PR domain peptides of rat RIZ (a.a. positions 36 to 151; SEQ ID NO: 2) and human RIZ (a.a. positions 37 to 152; see Figure 13, SEQ ID NO: 99) are identical except that the human RIZ contains a lysine at a.a. position 70, whereas the rat RIZ contains an arginine at the equivalent position (a.a. position 69). Additionally, analysis of a cDNA encoding a portion of the mouse RIZ

protein that includes blocks B and C of a PR domain peptide (Figure 14; SEQ ID NOS: 104 and 105) revealed that the deduced amino acid sequence (i.e. the first 75 amino acid in Figure 14 (SEQ ID NO: 105) is identical to the corresponding region of the PR domain in human RIZ (i.e. amino acids designated as positions 45-120 for human RIZ in Figure 13; SEQ ID NO:99).

The homology among PR domain peptides is evident from inspection of the amino acid sequences shown in Figure 13 (SEQ ID NOS: 98-101). If conserved amino acid substitutions are considered, the mammalian PR domain peptides (SEQ ID NOS: 98-100) are greater than about 33% homologous to each other over their entire length as shown. Moreover, the highly conserved blocks A, B and C are about 75%, 50% and 55% homologous, respectively, among the three mammalian proteins. In addition, the sequence designated X101, which lies between blocks B and C, is about 38% homologous among the mammalian PR domain peptides shown.

A PR domain peptide can be generally defined by the amino acid sequence, Y-A-X100-B-X101-C-Z, wherein Y is about 8 to 13 independently selected amino acids; X100 is about 20 to 35 independently selected amino acids; X101 is about 20 to 35 independently selected amino acids; Z is about 8 independently selected amino acids; A is Ile-X2-X3-Gly-X4-X5-X6-Gly-X7-X8-X9-X10, wherein X6 is Phe or Ile, X7 is Pro or Val, X10 is Gly or Lys, and X2, X3, X4, X5, X8 and X9 each is one independently selected amino acid, and, preferably, wherein X3 is Lys or Val, X5

is Arg or Lys, X6 is Phe, X7 is Pro, X9 is Ile or Val, and X10 is Gly; B is Ile-X11-X12-X13-X14-X15-X16-X17-X18-X19-X20-X21, wherein X11 is Asp or Cys, X15 is Pro or Glu, X16 is Glu or Asp, X20 is Trp or Phe, X21 is Met or Leu, and X12, X13, X14, X17, X18 and X19 each is an independently selected amino acid, and, preferably, X11 is Asp, X12 is Gly or Ala, X17 is Lys or Val, X18 is Ser or Gly, X19 is Asn or Ser, and X20 is Trp, and wherein C is X22-X23-L-X24-X25-X26-X27-X28-X29-X30-D, wherein X22 is Glu or Gln, X23 is Glu or Asn, X24 is Leu or Ile, X25 is Val or Leu, X30 is Arg or Glu, and X27 and X29 each is an independently selected amino acid, and X26 and X28 each independently is absent or is one amino acid; and, preferably, X23 is Glu, X24 is Leu, X26 is Trp or Phe, X27 is Tyr or Met, and X28, when present, is Asn or Lys.

The present invention provides a PR domain as a component of a transcriptional regulator by operably linking the domains to a peptide that binds to a particular gene promotor or enhancer, wherein binding of the fusion protein to the target gene can alter expression of the target gene. As used herein, the term "transcriptional regulation" includes transcriptional repression and transcriptional activation. Particularly preferred transcriptional regulators containing a PR domain include amino acids 17 to 900 of SEQ ID NO: 2 or 18 to 910 of SEQ ID NO: 4.

The present invention further provides the PR domain or RIZ active fragment containing a PR domain as a probe to identify transcription factors or oncogenic

proteins in a cell that bind the PR domain. Methods to detect the interaction between a peptide and another cellular molecule are useful to detect binding between a PR domain and a transcriptional factor or oncogenic protein. Such methods are well known in the art and include, for example, the yeast two hybrid system (Fields and Song, Nature 340:245-246 (1989); Vojtek et al., Cell 74:205-214 (1993); Durfee et al., Genes Devel. 7:555-569 (1993), each of which is incorporated herein by reference). An *in vivo* transcription activation assay such as the yeast two hybrid system is particularly useful for identifying and manipulating the association of proteins. The results observed using such an assay likely mirror the interactions that naturally occur in a cell. Also, an *in vitro* assay utilizing, for example, a PR domain fused to GST (see Example II) can provide a simple, rapid and inexpensive method for identifying and isolating an PR domain binding target. Such an *in vitro* assay is particularly useful for confirming results obtained *in vivo*.

Human RIZ was cloned from human cDNA and genomic DNA libraries using the rat RIZ cDNA as a hybridization probe. A complete human RIZ cDNA sequence is shown in Figure 9A (SEQ ID NO: 3), which encodes a polypeptide having 1719 amino acid residues (see Figure 9A; SEQ ID NO: 4). Clones encoding two alternative 5'-termini of human RIZ were also obtained (see Figure 9B; SEQ ID NOS: 95 and 96).

The human RIZ gene coding region is encoded by eight exons and is located on chromosome 1p36 (see Example VI). Allelic variants of the human RIZ gene are disclosed herein. The RIZ D283 allele contains an Asp residue at a.a. position 283 (SEQ ID NO: 4) while the RIZ E283 allele contains an Glu (See Figure 9A for the D283 allele). The RIZ D283 allele is estimated to occur two times more frequently in the human population than the RIZ E283 allele. The difference between the two alleles is a T at nucleotide position 969 in the RIZ D283 allele (SEQ ID NO: 3) versus an A at the same position in the RIZ E283 allele.

The nucleotides encoding residue 283 of human RIZ (SEQ ID NO: 4) are contained within a region of triplet repeat nucleotides at nucleotide position 952-981 shown in Figure 9A (SEQ ID NO: 3). The triplet region encoding the D283 allele is $(GAA)_5(GAT)_5$ (designated 5-5) while the E283 allele is $(GAA)_6(GAT)_4$ (designated 6-4). Other alleles or mutations include the $(GAA)_4(GAT)_4$ triplet sequence (designated 4-4) that was detected in one allele of the Malme 3M melanoma cell line and the $(GAA)_7(GAT)_4$ sequence (designated 7-4) that was detected in one allele of SK-MEl-64 and MeWo melanoma cell lines (see Example VI).

The deduced rat (SEQ ID NO: 2) and human (SEQ ID NO: 4) RIZ amino acid sequences are 84% homologous. The rat and human RIZ proteins have similar sequence motifs, including cr1, cr2, ce1, PR, zinc finger, SH3,

SH2 and a nuclear localization signal, and are similar in size; rat RIZ (SEQ ID NO: 2) contains 1706 amino acids and has a calculated molecular weight of 187,437 Daltons while a human RIZ (SEQ ID NO: 4) contains 1719 amino acids and has a calculated molecular weight of 188,894 Daltons. A rabbit antiserum produced against rat RIZ (see Example II) cross reacts with human RIZ.

RIZ protein is present primarily in the cell nucleus. RIZ mRNA is expressed primarily in cells of neuroendocrine origin and is expressed in greater amounts in the fetus than in the adult (see Figure 11). RIZ is expressed in rat cells as a 250 kD phosphoprotein.

As used herein, the term "RIZ" means a protein having substantially the amino acid sequence of human RIZ as shown in Figure 9A (SEQ ID NO: 4) or of rat RIZ as shown in Figure 1A (SEQ ID NO: 2). The term "RIZ" is meant to include normal variants such as the 5'-terminus variants of RIZ (see Figure 9B; SEQ ID NOS: 95 and 96, 102 and 103.) and the allelic variants disclosed herein. Such normal variants can differ in amino acid sequence but share the same or similar functional activities such as binding to GTP, DNA or Rb (see Examples). A RIZ is referred to as a "normal RIZ" or a "wild-type RIZ", all of which are distinct from a mutant RIZ. In addition to the allelic variants, RIZ also can be a truncated RIZ protein encoded by a subset of the RIZ exons and that functions like a RIZ. Such a variant RIZ can be generated in the cell by alternative RNA splicing, which may be responsible for the three 5'-terminus variants of

human RIZ (see Figure 9A and 9B; SEQ ID NOS: 3, 4, 95, 96, 102 and 103).

The term "RIZ" also includes peptide fragments of a RIZ, including active fragments of a RIZ. As used herein, the term "active fragment" means a peptide portion of a full length RIZ protein that has at least one activity that is characteristic of the corresponding full length protein. A peptide portion of a rat RIZ having the sequence EIRCEEKPEDL (SEQ ID NO: 6) or a peptide portion of a human RIZ having the sequence EIRCDEKPEDL (SEQ ID NO: 91) are examples of active fragments of a RIZ that can bind to Rb. In addition, peptide portions of a RIZ containing a PR domain, for example, a.a. positions 36 to 151 of rat RIZ (SEQ ID NO: 2) or a.a. positions 37 to 152 of human RIZ (SEQ ID NO: 4; see, also, SEQ ID NO: 99), are examples of an active fragment of a RIZ protein, which may interact with transcription factors and can be involved in regulating transcription. In some cases, an active fragment of a RIZ protein is active only when present in the nucleus. In such a case, the active fragment can contain a nuclear localization signal such as that contained at a.a. positions 880 to 884 of human RIZ (SEQ ID NO: 4) or at positions 867 to 874 of rat RIZ (SEQ ID NO: 2). The nuclear localization signal is another example of an active fragment of a RIZ. Other RIZ activities that can be associated with an active fragment of a RIZ include the ability to bind DNA in a zinc ion-dependent manner,

the ability to bind GTP or an anti-RIZ antibody, or the ability to act as a hapten or immunogen to obtain an anti-RIZ antibody.

The present invention provides active fragments of a RIZ, such as EIRCEEKPEDL (SEQ ID NO: 6), or EIRCDEKPEDL (SEQ ID NO: 91), which contain substantially the amino acid sequence of the RIZ cr2 core motif, where the cysteine residue is required when the activity of the fragment is Rb binding. The cr2 core-motif containing fragments of human RIZ or rat RIZ are examples of active fragments of a RIZ. Such active fragments can be produced by recombinant DNA methods, by peptide synthesis or by enzymatic cleavage of a RIZ protein. The present invention also provides a non-naturally occurring polypeptide having incorporated therein a RIZ cr2 core motif. Such a polypeptide can be produced, for example, using well known recombinant DNA methods or by peptide synthesis.

A RIZ protein or a RIZ polypeptide containing a cr2 sequence such as the amino acid sequences EIRCEEKPEDL (SEQ ID NO: 6) and EIRCDEKPEDL (SEQ ID NO: 91) can bind to Rb and, therefore, is useful for isolating Rb from a sample. Purified Rb can be used, for example, as a control target in a diagnostic test to detect whether a subject has a mutated Rb. Additionally, Rb can be used to as a reagent to detect whether a sample has a RIZ which can bind to Rb or a mutant RIZ that fails to bind Rb. Mutations that affect the function of Rb and are diagnostic for cancer are well known in the art (see, for

example, Lee et al., In *Tumor Suppressor Genes*, Chapter 11, Marcell Decker (1990).

To purify Rb, RIZ protein can be contacted with the Rb containing sample under suitable conditions, which allow formation of a RIZ-Rb complex. Suitable conditions for complex formation can be determined empirically and include, for example, an appropriate buffer concentration and pH and time and temperature of incubation that permit binding of the RIZ to Rb. The RIZ-Rb complex can be separated from unbound material in the sample and Rb can be dissociated from the complex and obtained in substantially purified form.

Substantially purified Rb can be obtained, for example, by using affinity chromatography, in which a RIZ is bound to a solid support, the sample is applied to the support to allow binding of Rb to the RIZ, the support is washed to remove unbound material and Rb is eluted from the support. Useful solid supports include, for example, agarose, Sepharose™ or plastic. RIZ can be attached to a solid support by direct chemical coupling or by an indirect means such as an affinity interaction with an anti-RIZ antibody bound to the support. Other indirect means for coupling a RIZ to a support include incorporating one entity of a known ligand/receptor pair into the RIZ, with the corresponding entity coupled directly to the support. For example, biotin can be coupled to RIZ and avidin can be coupled directly to a solid support to bind the RIZ to the support. Also, RIZ can be expressed as a fusion to glutathionine S-

transferase (see Example II) and the fusion protein can be bound to a glutathionine coupled support.

The present invention also provides a RIZ binding reagent. As used herein the phrase "RIZ binding reagent" means a chemical or biological molecule that specifically binds to a RIZ. As used herein with reference to a RIZ, the term "specifically binds" means that under a defined set of conditions, the RIZ binding reagent interacts with a RIZ but not with an unrelated molecule or with a mutant RIZ. Rb and anti-RIZ antibody are examples of a RIZ binding reagent.

The invention also provides a mutant RIZ binding reagent. As used herein, the phrase "mutant RIZ binding reagent" means a chemical or biological molecule that specifically binds to a mutant RIZ but not to a wild-type RIZ. In this case, the mutant RIZ binding reagent, under a defined set of conditions, interacts with the mutant RIZ but not with a wild-type RIZ.

Rb and an antibody specific for a RIZ are examples of reagents that can specifically bind to a RIZ. As used herein, the term "antibody" is used in its broadest sense to include polyclonal and monoclonal antibodies, as well as polypeptide fragments of antibodies that retain a specific binding activity for a specific antigen of at least about 1×10^5 M^{-1}. One skilled in the art would know that a fragment such as Fab, $F(ab')_2$, Fv and Fd fragments of an anti-RIZ antibody, for example, can retain specific binding activity for a

RIZ and, thus, is included within the definition of an antibody. In addition, the term "antibody" as used herein includes naturally occurring antibodies as well as non-naturally occurring antibodies and fragments of antibodies that retain binding activity. Such non-naturally occurring antibodies can be constructed using solid phase peptide synthesis, can be produced recombinantly or can be obtained, for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains as described by Huse et al., Science 246:1275-1281 (1989), which is incorporated herein by reference.

An antibody specific for a RIZ can be prepared using well known methods as described, for example, by Harlow and Lane, *Antibodies: A laboratory manual* (Cold Spring Harbor Laboratory Press, 1988), which is incorporated herein by reference. For example, RIZ protein or a portion of the RIZ protein can be used as an immunogen, which can be prepared from natural sources or produced recombinantly or, in the case of a portion of the RIZ protein, can be chemically synthesized. Non-immunogenic peptides of RIZ protein can be made immunogenic by coupling to a carrier molecule such as bovine serum albumin or keyhole limpet hemocyanin as described, for example, by Harlow and Lane, *supra*, 1988. In addition, a RIZ fusion protein can be expressed as described in Example II. Such a fusion protein can be readily purified and used as an immunogen (see Example II). These methods can be used to produce various anti-RIZ antibodies.

Polyclonal antibodies can be raised, for example, in rabbits or goats. In addition, monoclonal antibodies can be obtained using well known methods (see, for example, Reed et al., Anal. Biochem. 205:70-76 (1992)), which is incorporated herein by reference; see, also, Harlow and Lane, *supra*, 1988). For example, spleen cells from a RIZ immunized mouse can be fused to an appropriate myeloma cell line such as SP2/0 or P3x653.Ag8 myeloma cells to produce hybridoma cells. Cloned hybridoma cell lines can be screened using a labeled RIZ immunogen to identify clones that secrete monoclonal antibodies. Hybridomas that express antibodies having a desirable specificity and affinity can be isolated and utilized as a continuous source of antibodies. A dependable source of monoclonal antibodies is desirable, for example, for preparing diagnostic kits as described below.

An antibody specific for a mutant RIZ protein also can be prepared using the above methods by immunizing with either the full-length mutant RIZ protein or with a fragment of the protein containing the mutation. Methods to direct the immune response to the mutant sequence also are well known in the art and include, for example, use of particular adjuvants or pre-prior tolerization of the animal to the wild-type RIZ sequence. Such tolerization can be performed by immunizing the animal with the wild-type RIZ in conjunction with administration of anti-T cell antibodies or immunosuppressive drugs. A monoclonal antibody to the mutant sequence can be obtained by screening a population

of hybridomas for those that express an antibody that binds the mutant RIZ sequence but not a wild-type RIZ sequence.

The invention also provides a substantially purified nucleic acid molecule, which encodes a RIZ such as a mammalian RIZ. For example, the invention provides substantially purified nucleic acid molecules having substantially the nucleotide sequences encoding human RIZ (Figure 9A; SEQ ID NO: 3) and rat RIZ (Figure 1A; SEQ ID NO: 1), including nucleotide sequences having alternative 5'-nucleotide sequences for human RIZ (Figure 9B; SEQ ID NOS: 95 and 96) and rat RIZ (Figure 1B; SEQ ID NO: 94). The cDNA sequence encoding human RIZ shown in Figure 9A (SEQ ID NO: 3) includes further 3' untranslated sequence, a portion of which is shown in Figure 9C (SEQ ID NO: 97). In addition, the invention provides substantially purified nucleic acid molecules encoding mouse RIZ that incldue substantially the nucleotide sequence shown in Figure 14 (SEQ ID NO: 104).

As used herein, the term "substantially purified nucleic acid molecule" means a nucleic acid molecule that is in a form that is relatively free from contaminating lipids, proteins, nucleic acids or other cellular material normally associated with a nucleic acid molecule in a cell. A substantially purified nucleic acid molecule can be obtained, for example, by recombinant DNA methods as described herein (see, also, Sambrook et al., *Molecular Cloning: A laboratory manual* (Cold Spring Harbor Laboratory Press 1989), which is

incorporated herein by reference) or can be chemically synthesized.

As used herein with reference to a RIZ, the term "substantially the nucleotide sequence" means, for example, the disclosed nucleotide sequences for human RIZ (SEQ ID NOS: 3, 95 and 96), as well as similar sequences that contain, for example, different nucleotides than shown in SEQ ID NOS: 3, 95 and 96, but that, as a result of the degeneracy of the genetic code, encode the same amino acid sequence as shown in SEQ ID NOS: 4, 102 and 103, respectively. In addition, the rat RIZ nucleotide sequences (SEQ ID NOS: 1 and 94) and the mouse RIZ nucleotide sequence (SEQ ID NO: 104) that corresponds to the rat RIZ sequence are considered to be substantially similar to the nucleotide sequence encoding human RIZ (SEQ ID NO: 3). For convenience, the coding strand for a nucleic acid molecule encoding a RIZ is shown. It should be recognized, however, that the complementary strand also is encompassed within the disclosed nucleic acid molecules. Thus, unless otherwise indicated, reference herein to a nucleic acid molecule or to a nucleotide sequence is meant to include the complementary sequence.

A nucleic acid molecule of the invention can encode a variant RIZ such as the allelic RIZ variants disclosed herein as well as variants of a RIZ that contain only particular exons of the gene that can be produced in a cell by alternative RNA splicing. In addition, a nucleic acid molecule of the invention can encode a portion of a RIZ such as an active fragment of a

RIZ containing the polypeptide EIRCEEKPEDL (SEQ ID NO: 6) and EIRCDEKPEDL (SEQ ID NO: 91), which binds to the Rb pocket or a RIZ fragment containig a PR domain peptide (see, for example, Figure 13; SEQ ID NOS: 98-101), which may bind to a transcription factor and be involved in transcriptional regulation.

The invention also provides a nucleotide sequence that specifically hybridizes to a portion of a nucleic acid molecule encoding a mammalian RIZ under relatively stringent hybridization conditions. As used herein with reference to a RIZ, the term "specifically hybridizes" means that under a defined set of hybridization conditions, the nucleotide sequence can interact with a RIZ encoding nucleic acid molecule but not with an unrelated nucleic acid molecule. A nucleotide sequence that specifically hybridizes to a RIZ can be complementary to a nucleotide sequence encoding a RIZ or can be a RIZ coding sequence or a portion thereof.

A nucleotide sequence that specifically hybridizes to a nucleic acid molecule encoding a RIZ or a mutant nucleic acid molecule encoding a RIZ should be at least ten nucleotides in length and can be prepared, for example, by restriction endonuclease digestion of a cloned nucleic acid molecule encoding a RIZ or by PCR amplification of a portion of the nucleic acid molecule shown in Figure 1A or 1B (SEQ ID NOS: 1 and 94) or Figure 9A, 9B or 9C (SEQ ID NOS: 3, 95, 96 and 97), or by chemical synthesis. A nucleotide sequence that can hybridize to one or more of the nucleotide sequences

encoding the highly conserved block A, B or C of a RIZ PR domain peptide (see Figure 13) can be particularly useful, for example, to identify nucleic acid molecules that encode other members of PR domain peptide-containing family of proteins.

Relatively stringent hybridization conditions can be determined empirically or can be estimated based, for example, on the relative GC:AT content of the hybridizing nucleotide sequence and the target sequence, the length of the hybridizing nucleotide sequence and the number, if any, of mismatches between the hybridizing nucleotide sequence and the target sequence. If desired, a hybridizing nucleotide sequence can be detectably labeled and used as a probe or can be used as a primer for PCR. Methods for detectably labeling a nucleotide sequence are well known in the art (see, for example, Sambrook et al., *supra*, 1989; see, also, Ausubel et al., *Current Protocols in Molecular Biology* vol. 2, chapter 10 (Greene Publ., NY 1989), which is incorporated herein by reference).

As used herein, the term "mutant nucleic acid encoding a RIZ" includes nucleic acid molecules having a mutation in an exon, thus encoding a mutant RIZ protein, as well as nucleic acid molecules having a mutation in a region of the RIZ gene other than the exons. A mutation in the RIZ gene occurring outside the exons can involve a regulatory element of the gene that modulates the expression of the RIZ in a cell. Such regulatory elements that can be mutated include, for example, the

promoter, enhancer, ribosomal binding site or intron-exon splice junctions. The term "mutant RIZ" also includes peptides of a mutant RIZ, including active fragments of a mutant RIZ.

A mutation that occurs in a regulatory element of the RIZ gene can have a significant impact on the level of expression of a RIZ in a cell. In addition, a mutation in a RIZ exon that codes for a stop codon within the reading frame of the RIZ can produce a truncated RIZ that may be inactive, have an altered activity or be subject to rapid proteolysis in the cell. Similarly, a deletion involving a substantial portion of the gene encoding the RIZ can result in a loss of RIZ expression.

As used herein, the term "mutant RIZ" includes any RIZ having a mutation in a RIZ exon that results in the expression of a RIZ having a functional activity differing from that of a wild-type RIZ normally expressed by a cell. A change in a functional activity characteristic of a mutant RIZ can result from one or more amino acid additions, deletions or substitutions in the wild-type RIZ sequence. Such mutations can arise spontaneously or can be resident in the population and inherited from generation to generation as occurs, for example, with Rb. A mutant RIZ can have a change in the nucleotide at position 437 in human RIZ (SEQ ID NO: 3)from a G to an A, which results in the expression of mutant RIZ having a Tyr residue instead of a Cys residue at a.a. position 106 (SEQ ID NO: 4).

The present invention also provides a nucleotide sequence that specifically hybridizes to a mutant nucleic acid molecule encoding a RIZ under relatively stringent conditions but not to a wild-type RIZ. In this case, the hybridizing sequence should be complementary to a portion of the RIZ gene containing the mutation.

The expression of a particular RIZ allele can be altered in a cancer cell due to a mutation in the RIZ gene. As disclosed herein, some melanoma tumor cells fail to express mRNA encoding one of two RIZ gene alleles present in the cells (see Example VI). The unexpressed allele likely contains a mutation outside the RIZ coding sequence that affects RIZ expression. Detection of such mutations through the RIZ protein or the RIZ gene can be diagnostic of a pathology such as a cancer.

A mutant RIZ can be obtained, for example, by site directed mutagenesis of a nucleic acid molecule encoding a RIZ, then screening the mutagenized nucleic acid molecule to identify an encoded mutant RIZ. Mutations that affect a functional activity of a RIZ such as Rb binding, DNA binding or GTP binding can be detected by screening for mutants that have lost such activities. Expression in a cell of a mutant RIZ such as mutant human RIZ, which can bind Rb, for example, but lacks another RIZ activity, can alter the association of wild type RIZ with Rb and can affect a function of a cell such as the ability of the cell to proliferate.

The ability of a RIZ to be expressed in the nucleus together with its ability to bind DNA, Rb and GTP (see Example II and IV) and its homology with Blimp-1 (PRD1-BF1) differentiation factor indicates that RIZ can function as a transcriptional regulatory protein or cell differentiation factor. Thus, a function of a cell can be modulated by expressing a RIZ in a cell, where the expressed RIZ can bind to Rb and to DNA in the cell. Cell function can also be modulated through the ability of the complete RIZ protein or an active fragment of RIZ containing the PR domain peptide to act as a transcriptional regulator (see Example VII).

As used herein, the term, "a function of a cell" means a cell activity, including, for example, proliferation and differentiation. As used herein, the term "modulate" means increase or decrease. As disclosed herein, the function of a cell can be modulated due to an altered level of expression of a RIZ or expression of a mutant RIZ in a cell.

The present invention provides methods for modulating a function of a cell by expressing in the cell a DNA sequence encoding a RIZ or an active fragment of a RIZ. Such a DNA sequence can be expressed by introducing into a host cell an appropriate expression vector having gene regulatory elements operably linked with the RIZ encoding nucleotide sequence. The expression vector can provide constitutive expression of the polypeptide or, if desired, inducible expression. Expression vectors having the appropriate gene regulatory elements can be purchased

from commercial sources or can be constructed using well known methods. For therapeutic purposes, cells can be transfected in tissue culture, then administered to a subject, or a viral vector can be used to introduce a RIZ encoding nucleic acid into a cell in a subject.

Because the transcriptional activity of RIZ requires localization of RIZ to the cell nucleus, active fragments of RIZ can, depending on the size of the fragment, require inclusion of a nuclear localization signal. For example, a RIZ active fragment can include the nuclear localization signal endogenous to rat RIZ at a.a. position 867-874 (SEQ ID NO: 2) or human RIZ at a.a. position 880-884 (SEQ ID NO: 4). Alternatively, the nuclear localization signal included in the RIZ active fragment can be derived from a non-RIZ protein. A variety of nuclear localization sequences are known in the art that can direct proteins to the cell nucleus (see for example, Dingwall et al. EMBO J. 8:69-71 (1987), which is incorporated herein by reference).

As disclosed herein, RIZ can regulate the growth of normal adult cardiac muscle cells and prevent proliferation of surviving cells following cardiac muscle cell death. RIZ can function to maintain cells in the G_1 phase of the cell cycle by interacting with Rb through the cr2 domain of RIZ. In addition, the functional differentiation state of a cell, which involves maintenance of a cell in the G_1 phase of the cell cycle

can be affected through the action of other RIZ domains such as the GTP binding domain and the zinc finger domains.

The regeneration of cardiac muscle cells can be promoted in a subject with cardiac damage by directly decreasing the activity of a RIZ or by decreasing the activity of Rb that occurs subsequent to RIZ binding. The activity of a RIZ can be decreased in such cells by introducing into the cells an expression vector having an expression control sequence operatively linked to a nucleotide sequence encoding a mutant RIZ polypeptide or an active fragment that can bind to Rb but lacks the growth suppressing properties of RIZ. The sequences EIRCEEKPEDL (SEQ ID NO: 6) and EIRCDEKPEDL (SEQ ID NO: 91) are examples of such a peptide.

As used herein, the term "growth suppressing properties of RIZ" means the ability of RIZ to effect the differentiation and the maintenance of cells in G_1. In fact, the cell may be in an extended G_1 phase or a G_0 phase or may be blocked at the G_0/G_1 boundary. For convenience, any such cells are referred to as being maintained or suppressed in G_1. The growth suppressing or differentiating properties of a RIZ can be mediated by regions of the molecule outside the cr2 domain or in conjunction with cr2 that is involved in binding to the Rb pocket.

Neurons, like myocardial cells, normally do not proliferate in the adult. RIZ is preferentially

expressed in neural cells (see Example V), indicating a role for RIZ in mediating G_1 suppression and differentiation of these cells. The ability to induce proliferation in neural cells can be useful for healing after injury of neural tissue treating neurodegenerative diseases such as Parkinson's disease, Huntington's disease or Alzheimer's disease or paralysis or motor neuron disorders. Thus, the disclosed methods for decreasing the activity of a RIZ protein in a muscle cel similarly can provide a therapy for a neurodegenerative disease.

As disclosed herein, the RIZ gene is in chromosome band 1p36; therefore, mutations in the nucleotide sequence encoding a RIZ can be involved in the development of cancer, particularly cancers such as melanoma, neuroblastoma, leukemia, and breast cancers known to be associated with deletions or rearrangements involving 1p36. Melanoma cells can show a reduction or absence of expression of a RIZ allele and, a reduction i the overall amout of RIZ protein expressed in the cell. These results indicate that melanoma can be characterize by a reduced level of RIZ protein and, thus, a reduced level of RIZ function, which may explain the selective growth advantage of melanoma tumor cells that occurs following alterations in distal chromosomal 1p (Dracopol et al., Proc. Natl. Acad. Sci., *USA* 86:4614-4618 (1989)) Mutations in distal chromosome 1p that affect the level of expression of RIZ may be responsible for the increase risk of melanoma observed in survivors of heritable retinoblastoma, which occurs without homozygous

inactivation of the Rb gene. Since Rb binds to RIZ, a decreased level of Rb-RIZ complex in a tumor cell, resulting from a reduced expression of a RIZ allele, can result in a loss in Rb tumor suppressor activity in the cell without Rb mutation.

Further support for the loss of RIZ function and the development of cancer is provided by the disclosure that RIZ is a differentiation factor. As such, a mutant RIZ can affect the regulation of cell growth by binding to the Rb pocket, a site in the Rb molecule that is involved in regulating cell proliferation. Thus, the present invention provides methods for restoring normal cell growth to a cancer cell that has a mutated or missing RIZ allele by expressing a normal RIZ protein in the cell. As shown in Example VI, increasing the expression of RIZ in a tumor cell by transfecting the cell with a RIZ expresssion vector results in decreased cell growth.

The disclosure that RIZ can modulate a function of a cell by binding to a second molecule such as Rb or a nucleic acid such as DNA or RNA provides a means to identify agents that can effectively alter the association of a RIZ with a second molecule in a cell and, as a result, modulate a function of a cell. Thus, the present invention provides a screening assay useful for identifying an effective agent, which can alter the association of a RIZ with a second molecule. A second

molecule that binds to a RIZ can also be a transcription factor or an oncogenic protein. Such second molecules can bind to the PR domain of RIZ.

An effective agent that can decrease the association of a RIZ with a second molecule such as Rb or that can decrease the activity of a RIZ can be useful for releasing a cell from Rb-mediated G_1 arrest. Alternatively, an effective agent that increases the association of a RIZ with a second molecule such as Rb or DNA or increases the activity of a RIZ can be useful for reducing the unrestricted growth of a cancer cell by providing a stronger G_1 arrest signal in the cell.

A nucleotide sequence that specifically binds to a RIZ can be detected by using methods well known in the art (see for example, El-Deiry et al., Nat. Genet. 1:45 (1992), which is incorporated herein by reference). Genomic DNA can be processed, for example, by sonication to produce uniform-sized fragments, which can be screened for the ability to bind to a RIZ. Genomic DNA sequences that bind to a RIZ can be isolated using, for example, an anti-RIZ antibody and Protein A affinity chromatography. The isolated DNA sequences can be amplified by PCR, which can be facilitated by ligating the original genomic DNA fragments to "catch linkers" (El-Deiry et al., *supra*, 1992) suitable for annealing to PCR primers.

Random oligonucleotides consisting of at least about ten nucleotides and including "catch linkers" also can be screened to identify sequences that can bind a

RIZ. For example, RIZ protein can be immobilized to a filter, then incubated with the oligonucleotides under conditions that allow the RIZ to bind relatively specifically to a RIZ binding sequence. Unbound oligonucleotides can be washed from the filter, then specifically bound sequences can be eluted and amplified by PCR. Following three or more cycles of binding, elution and amplification, a consensus RIZ binding sequence can be obtained. If desired, the consensus RIZ binding sequence can be used to screen a genomic DNA library to obtain genomic DNA sequences containing the RIZ binding sequence.

An agent can be a chemical or biological molecule such as a simple or complex organic molecule, a peptide, a peptido-mimetic, a protein, a carbohydrate or an oligonucleotide that has the potential for altering the association of a RIZ with a second molecule or altering an activity of a RIZ. With reference to a RIZ, the term "effective agent" means an agent that can, in fact, alter the association of RIZ with a second molecule or can alter the activity of a RIZ.

An effective agent can be, for example, a nucleic acid molecule that encodes a RIZ or a mutant RIZ or is complementary to a RIZ- or mutant RIZ-encoding nucleotide sequence. Such nucleic acid molecules can be contained within an expression vector having the RIZ encoding sequence operably linked to an expression control sequence. An effective agent also can be an antisense RIZ or a ribozyme complementary to a RIZ mRNA

sequence. Such agents can reduce the level of expression of a RIZ in a cell and, as a consequence, can alter the amount of a RIZ that is associated with a second molecule in a cell.

As used herein with reference to a RIZ, the term "alter the association" means that the association of a RIZ and a second molecule either is increased or is decreased due to the presence of an effective agent. As a result of an altered association of RIZ with a second molecule in a cell, the activity of the RIZ or second molecule can be increased or decreased, which can modulate a function of a cell. As used herein with reference to a RIZ, the term "alter the activity" means that the effective agent can increase or decrease the activity of RIZ in a cell, such as by altering the association of a RIZ with the second molecule as described above by modifying, for example, an activity of a RIZ that occurs consequent to binding a second molecule.

An effective agent that alters the association of a RIZ with a second molecule can interfere with the ability of the RIZ and the second molecule to associate or can cause the dissociation of a bound RIZ-second molecule complex. In the presence of an effective agent, the association of a RIZ with a second molecule can be altered, which can alter the activity of the RIZ or the second molecule in the cell. As a result of the altered activity, a cell function such as the ability of a cell to proliferate can be modulated. Thus, the

identification of an effective agent that alters the association of a RIZ with a second molecule provides a means to modulate cell proliferation.

An effective agent that alters the association of a RIZ and Rb can be useful as a medicament to treat a pathology characterized, in part, by excessive cell growth such as occurs in a cancer or by insufficient cell growth such as occurs in a tissue that fails to regenerate in response to cell death. A peptide having the sequence EIRCEEKPEDL (SEQ ID NO: 6) or EIRCDEKPEDL (SEQ ID NO: 91), which contains the cr2 core motif of RIZ, is an example of an effective agent. Either of the peptides can alter the association between a RIZ and Rb (see Example II) and can induce cells such as adult cardiac muscle cells or adult neural cells to proliferate, which can regenerate heart function or neural function, respectively, following injury or disease.

The present invention also provides *in vitro* screening assays to detect an effective agent. Such screening assays are particularly useful in that they can be automated, which allows for high through-put screening, for example, of randomly or rationally designed agents such as drugs, peptido-mimetics or peptides in order to identify agents that effectively alter the association of a RIZ and a second molecule or modulate a function of a cell.

An *in vitro* screening assay can utilize, for example, RIZ or a RIZ fusion protein such as a glutathione-S-transferase-RIZ fusion protein (GST-RIZ; see Example II). For *in vitro* screening assays, the RIZ or RIZ fusion protein can be attached to a solid substrate, provided the attached RIZ maintains the ability to associate with a particular second molecule. For example, when human RIZ is used in the assay, the solid substrate can contain a covalently attached anti-RIZ antibody to bind RIZ to the substrate (see Example II). Alternatively, a GST-RIZ fusion protein can be used in the assay and the solid substrate can contain covalently attached glutathione, which is bound by the GST component of the GST-RIZ fusion protein. Similarly, a second molecule or a GST-second molecule fusion protein can be used in an *in vitro* assay as described herein.

An *in vitro* screening assay can be performed by allowing, for example, a RIZ or RIZ-fusion protein to bind to the solid support, then adding a second molecule and an agent to be tested. Alternatively, a second molecule or a second molecule-fusion protein can be attached to the solid support and RIZ and an agent to be tested are added. Control reactions, which do not contain an agent, can be performed in parallel. Following incubation under suitable conditions, which include, for example, an appropriate buffer concentration and pH and time and temperature of incubation that permit binding of a RIZ and a second molecule, the amount of the

RIZ and second molecule that have associated in the absence of an agent and in the presence of an agent can be determined.

The association of a RIZ and a second molecule can be detected, for example, by attaching a detectable moiety such as a radionuclide or a fluorescent label to the second molecule and measuring the amount of label that is associated with the solid support, wherein the amount of label detected indicates the amount of association of the second molecule and RIZ. By comparing the amount of specific binding in the presence of an agent as compared to the control level of binding, an effective agent, which alters the association of a RIZ and a second molecule, can be identified. Such an assay is particularly useful for screening a panel of agents such as a peptide library in order to detect an effective agent.

In an *in vitro* screening assay as disclosed herein, the order in which the components are added can be informative. For example, the agent to be detected can be combined with a RIZ prior to adding a second molecule, can be combined with a second molecule prior to adding a RIZ or can be added after allowing binding of the RIZ and the second molecule. Depending on the relative affinities of the components in the reaction mixture for each other, the order of addition and the time between mixing the first two components and adding the remaining component can be manipulated to detect effective agents with varying properties.

The methods for identifying an effective agent that alters the association of RIZ with a second molecule, can be performed to determine, for example, whether the agent can dissociate a bound RIZ-second molecule complex. For this purpose, a RIZ is first contacted with a second molecule under conditions suitable for forming a RIZ-second molecule complex and thereafter the complex is contacted with the effective agent.

The invention also provides methods for identifying an effective agent that alters the association of a RIZ and a second molecule in a test sample containing the RIZ and the second molecule. As used herein, the term "test sample" means a cell or tissue specimen that is obtained from a subject and is to be examined for expression of RIZ protein or a nucleic acid molecule encoding RIZ. A test sample can be obtained, for example, during surgery or by needle biopsy. The test sample can be, for example, a soluble lysate of a cell preparation obtained by treating the cells with a solubilizing agent such as a non-ionic detergent.

A soluble lysate or other form of test sample can be examined by a gel-shift assay to determine the proportion of a RIZ and a second molecule that are associated as a complex. In this assay, the test sample is electrophoresed in a non-denaturing gel such as a low percentage polyacrylamide gel with a buffer containing 50 mM Tris (pH 8.5), 0.4 M glycine, 2 mM EDTA and

3% glycerol. By adjusting the buffer conditions, gel concentration or other parameters of electrophoresis well known in the art, electrophoretic separation of a free second molecule, a free RIZ and a second molecule-RIZ complex in the test sample can be achieved. After electrophoresis, the identity of proteins in the gel can be determined by immunoblotting using antibodies specific for the second molecule or the RIZ. Methods for performing immunoblotting using an enzyme or radioisotope labeled primary or secondary antibody are well known in the art (see, for example, Harlow and Lane, *supra*, 1988).

If desired a separate gel can be produced and immunoblotted with either anti-second molecule antibodies or anti-RIZ antibodies. Each gel can contain known amounts of both the second molecule and the RIZ to be detected to provide standards for quantitation and specificity of the blot. The amount of a second molecule-RIZ complex in a test sample treated with an agent suspected of being able to alter the association of the second molecule with RIZ can be compared to a control test sample not treated with the agent in order to identify an effective agent, which increases or decreases the proportion of the second molecule-RIZ complex in the treated relative to the control test sample.

The present invention provides methods to modulate a function of a cell by contacting the cell with an effective agent. As used herein, the term "contacting" means providing within sufficient proximity such that the effective agent can interact with a target.

Thus, an effective agent can be contacted with Rb *in vitro*, or can be contacted with a cell, provided the effective agent can enter the cell to interact with RIZ or a second molecule. For example, a small molecule effective agent can enter a cell passively such as through pores in the cell membrane or through the lipid bilayer of the cell. An effective agent also can enter a cell by active means such as through pinocytosis, endocytosis, phagocytosis or through an energy driven specific transport mechanism.

Methods for introducing and expressing a RIZ in a cell can be performed using well known expression vectors and gene transfer methods (for example, see Sambrook et al., *supra*, 1989; see, also, Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (W. H. Freeman and Co.; New York, 1990), which is incorporated herein by reference). Such gene transfer methods include, for example, transfection methods such as calcium phosphate precipitation, electroporation or lipofection, or viral infection. For convenience, the term "transfected cell" is meant to include any cell having an exogenous nucleic acid molecule introduced therein. Transfected cells useful for expressing large amounts of RIZ protein can be prokaryotic or eukaryotic and include, for example, bacterial cells such as *E. coli*, yeast cells, insect cells or mammalian cells such as COS cells or Chinese hamster ovary (CHO) cells.

An expression vector useful for expressing a RIZ or a mutant RIZ in a cell contains an expression

control sequence operatively linked to a nucleotide sequence encoding a RIZ. An expression control sequence that is operatively linked to a nucleic acid sequence can direct the transcription and translation of the nucleic acid sequence *in vitro* or in an appropriate host cell. Expression control elements are well known in the art and include, for example, promoters, enhancers and appropriate start and stop codons. In particular, a tissue specific expression control element can provide a means to selectively express a RIZ or mutant RIZ in a cell. Tissue specific control elements are well known in the art and include, for example, the muscle creatine kinase enhancer for restricting expression to muscle cells and the Purkinje cell protein-2 promoter for restricting expression to Purkinje cells (Vandaele et al., Genes Devel. 5:1136-1148 (1991), which is incorporated herein by reference).

Viral vectors that are compatible with a targeted cell are particularly useful for introducing a nucleic acid molecule encoding a RIZ or a mutant RIZ into a cell. For example, recombinant adenoviruses having general or tissue-specific promoters can be used to deliver a nucleic acid encoding RIZ into a variety of cell types in various tissues and can direct expression of the nucleic acid in the target cell (Lebkowski et al., U.S. Patent 5,354,678, issued October 11, 1994, which is incorporated herein by reference). Recombinant adeno-associated viruses also are useful for introducing a nucleic acid molecule encoding RIZ into a cell and have the added advantage that the recombinant virus can stably

integrate into the chromatin of even quiescent non-proliferating cells such as neurons of the central and peripheral nervous systems (Lebkowski et al., Mol. Cell. Biol. 8:3988-3996 (1988), which is incorporated herein by reference).

Such viral vectors are particularly useful where it is desirable to introduce a nucleic acid molecule encoding a RIZ or a mutant RIZ into a cell in a subject, for example, for gene therapy. Viruses are specialized infectious agents that can elude host defense mechanisms and can infect and propagate in specific cell types. The specificity of viral vectors for particular cell types can be utilized to target predetermined cell types. Thus, the selection of a viral vector will depend, in part, on the cell type to be targeted. For example, if a neurodegenerative disease is to be treated by decreasing the level of RIZ in neuronal cells affected by the disease, then a viral vector that targets neuronal cells can be used. A vector derived from a herpes simplex virus is an example of a viral vector that targets neuronal cells (Battleman et al., J. Neurosci. 13:941-951 (1993), which is incorporated herein by reference).

A viral vector that is specific for a particular blood cell or its precursor cell can be used to introduce a nucleic acid molecule encoding a RIZ or a mutant RIZ into a hematopoietic cell from a subject having a pathological condition of the hematopoietic system. A vector based on a human immunodeficiency virus

is an example of such a viral vector (Carroll et al., J. Cell. Biochem. 17E:241 (1993), which is incorporated herein by reference). In addition, a viral vector or other vector can be constructed to express a nucleic acid encoding a RIZ in a tissue specific manner by incorporating a tissue-specific promoter or enhancer into the vector (Dai et al., Proc. Natl. Acad. Sci. *USA* 89:10892-10895 (1992), which is incorporated herein by reference).

Retroviral vectors can be particularly useful for introducing a nucleic acid molecule encoding a RIZ or a mutant RIZ into a cell *in vivo*. Retroviral vectors can be constructed either to function as infectious particles or as non-infectious particles that undergo only a single initial round of infection. In the former case, the genome of the virus is modified so that it maintains the necessary genes, regulatory sequences and packaging signals to synthesize new viral proteins and RNA. However, genes conferring oncogenic potential of these viruses are destroyed or removed. After the viral proteins are synthesized, the host cell packages the RNA into new viral particles, which can undergo further rounds of infection. The viral genome also is engineered to encode and express the desired recombinant gene.

In the case of non-infectious viral vectors, a helper virus genome is required to provide the structural genes necessary to encode for the viral structural proteins. However, the helper virus is mutated to destroy the viral packaging signal required to

encapsulate the helper viral RNA into viral particles. Thus, only the recombinant viral vector containing the gene of interest and a functional packaging signal, but lacking viral structural genes can be incorporated into a virus particle. Although this new virus can infect a target cell, no further infectious virus can be produced since there are not viral structural genes provided. Methods for constructing and using viral vectors are known in the art and reviewed, for example, in Miller and Rosman, Biotechniques 7:980-990 (1992), which is incorporated herein by reference. The specific type of vector will depend upon the intended application. These vectors are well known and readily available within the art or can be constructed by one skilled in the art.

For gene therapy, a vector containing a nucleic acid encoding a RIZ or a mutant RIZ can be administered to a subject by various methods. For example, if viral vectors are used, administration can take advantage of the target specificity of the vectors. In such cases, there is no need to administer the vector locally at the diseased site. However, local administration can be a particularly effective method of administering a nucleic acid molecule. In addition, administration can be via intravenous or subcutaneous injection into the subject. Following injection, the viral vectors will circulate until they recognize host cells with the appropriate target specificity for infection. Injection of viral vectors into the spinal fluid also can be an effective mode of administration, for example, in treating a neurodegenerative disease.

Receptor-mediated DNA delivery approaches also can be used to deliver a nucleic acid molecule encoding a RIZ or a mutant RIZ into a cell in a tissue-specific manner using a tissue-specific ligand or an antibody that is non-covalently complexed with the nucleic acid molecule via a bridging molecule (Curiel et al., Hum. Gene Ther. 3:147-154 (1992); Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987), each of which is incorporated herein by reference). Direct injection of a naked or a nucleic acid molecule encapsulated, for example, in cationic liposomes also can be used for stable gene transfer into non-dividing or dividing cells *in vivo* (Ulmer et al., Science 259:1745-1748 (1993), which is incorporated herein by reference). In addition, a nucleic acid molecule encoding a RIZ can be transferred into a variety of tissues using the particle bombardment method (Williams et al., Proc. Natl. Acad. Sci. *USA* 88:2726-2730 (1991), which is incorporated herein by reference). Such nucleic acid molecules can be linked to the appropriate nucleotide sequences required for transcription and translation.

A particularly useful mode of administration of a nucleic acid encoding a RIZ or mutant RIZ is by direct inoculation locally at the site of the disease or pathological condition. Local administration can be advantageous because there is no dilution effect and, therefore, the likelihood that a majority of the targeted cells will be contacted with the nucleic acid molecule is increased. Thus, local inoculation can alleviate the targeting requirement necessary with other forms of

administration and, if desired, a vector that infects all cell types in the inoculated area can be used. If expression is desired in only a specific subset of cells within the inoculated area, then a promotor, an enhancer or other expression element specific for the desired subset of cells to be targeted can be linked to the nucleic acid molecule. Vectors containing such nucleic acid molecules and regulatory elements can be viral vectors, viral genomes, plasmids, phagemids and the like. Transfection vehicles such as liposomes also can be used to introduce a non-viral vector into recipient cells. Such vehicles are well known in the art.

An alternative method of modulating a function of a cell is to introduce a nucleic acid molecule having a nucleotide sequence encoding an antisense RIZ or a ribozyme specific for a RIZ mRNA into the cell. Such a nucleotide sequence is included within the meaning of an effective agent since it can alter the expression level of RIZ and thus alter the association of a RIZ with a second molecule.

An antisense RIZ or a ribozyme specific for a RIZ mRNA can be complementary to the nucleotide sequence of a RIZ such as the nucleotide sequence of Figure 1A (SEQ ID NO: 1), Figure 9A (SEQ ID NO: 3) or Figure 14 (SEQ ID NO: 104). An antisense RIZ or ribozyme specific for RIZ mRNA can be introduced into a cell using the methods and vectors described above. Chemically synthesized nucleotide sequences also can be administered directly to cells. Synthetic antisense or ribozyme

oligonucleotides can be prepared using well known methods or can be purchased from commercial sources and, if desired, can incorporate nucleotide analogs to increase the resistance of the oligonucleotide to degradation by nucleases in a cell. Synthetic antisense or ribozyme sequences can be active in a cell after contact with and uptake by the cell.

An effective agent can be administered *in vivo* as a pharmaceutical composition containing the effective agent and a pharmaceutically acceptable carrier. Pharmaceutically acceptable carriers are well known in the art and include, for example, aqueous solutions such as physiologically buffered saline or other solvents or vehicles such as glycols, glycerol, oils such as olive oil or injectable organic esters. Methods to formulate pharmaeuctical compositions are well known in the art (see, for example, Renaso et al. Remington Pharmaceutical Sciences, Mack Publishing Co., Eaton, Penn. (1990), which is incorporated herein by reference).

A pharmaceutically acceptable carrier can contain physiologically acceptable compounds that act, for example, to stabilize or to increase the absorption of an effective agent. Such physiologically acceptable compounds include, for example, carbohydrates, such as glucose, sucrose or dextrans, antioxidants, such as ascorbic acid or glutathione, chelating agents, low molecular weight proteins or other stabilizers or excipients. One skilled in the art would know that the choice of a pharmaceutically acceptable carrier,

including a physiologically acceptable compound, depends, for example, on the route of administration of the composition. One skilled in the art would know that a pharmaceutical composition containing an effective agent can be administered to a subject by various routes including, for example, by direct instillation, orally or parenterally, such as intravenously, intramuscularly, subcutaneously or intraperitoneally. The composition can be administered by injection or by intubation. The pharmaceutical composition also can be incorporated, if desired, into liposomes or microspheres or can be microencapsulated in other polymer matrices (Gregoriadis, Liposome Technology, Vol. 1 (CRC Press, Boca Raton, FL 1984), which is incorporated herein by reference). Liposomes, for example, which consist of phospholipids or other lipids, are nontoxic, physiologically acceptable and metabolizable carriers that are relatively simple to make and administer.

In order to modulate a function of a cell, an effective agent is administered in an effective amount, which can be determined using methods well known to those in the art (see, for example, Renaso et al. *supra*, 1990). As used herein, the term "effective amount" means the amount that produces a desired effect. Thus, an effective amount of an effective agent can alter the association of a RIZ and Rb in a cell and can have a functional effect on the ability of a target cell to increase or decrease its ability to enter the cell cycle. Administration of an effective amount of an effective agent *in vivo* can reduce symptoms associated with a

disease being treated.

The total effective amount can be administered to a subject as a single dose, either as a bolus or by infusion over a relatively short period of time, or can be administered using a fractionated treatment protocol, in which the multiple doses are administered over a more prolonged period of time. One skilled in the art would know that the concentration of an effective agent needed to obtain an effective amount in a subject depends on many factors including the age and general health of the subject as well as the route of administration and the number of treatments to be administered, as well as the chemical form of the effective agent (see, for example, Renaso et al. *supra*, 1990). In view of these factors, the skilled artisan would adjust the particular dose so as to obtain an effective amount for subject being treated.

The present invention also provides methods for detecting the presence of a RIZ in a test sample by detecting the RIZ protein or a nucleic acid molecule encoding RIZ. In addition, methods are disclosed for diagnosing a pathology that is characterized, in part, by an increased or decreased ability of a cell to enter the cell cycle by determining whether cell proliferation or lack thereof is due, for example, to increased or decreased expression of a RIZ or a mutant RIZ in the cell. The identification of such a pathology can allow for intervention therapy using an effective agent as described above.

A test sample can be obtained from a subject having a pathology characterized by increased or decreased cell function and can be compared to a control sample from a normal healthy subject to determine whether the cells in the test sample have an increased or decreased level of a RIZ or a mutant RIZ. The level of RIZ protein in a cell can be determined by contacting a sample with a RIZ binding reagent such as an anti-RIZ antibody or Rb. For example, the level of RIZ in a cell can be determined by well known immunoassay or immunohistochemical methods using an anti-RIZ antibody (see, for example, Reed et al., *supra*, 1992; see, also, Harlow and Lane, *supra*, 1988). In addition, the expression of a mutant RIZ can be detected, for example, by an antibody that specifically binds to the mutant RIZ but not to wild-type RIZ.

The detection of a RIZ by binding to an antibody and to Rb can provide complementary information. For example, the antibody can be used to determine the total level of RIZ expressed, while Rb binding can be used to determine the fraction of RIZ that is bound to Rb. Because Rb can bind to other proteins in a cell, it is advantageous to first isolate RIZ from a cell prior to detecting the fraction of RIZ that is bound to Rb.

An increased or decreased level of expression of a RIZ in a cell in a test sample can be determined by comparison to an expected normal level for the RIZ in a particular cell type. A normal range of RIZ levels in various cell types can be determined by sampling a

statistically significant number of normal cell types, which can be obtained from healthy subjects. In addition, a control sample can be evaluated in parallel with a test sample in order to determine whether a pathology characterized by increased or decreased cell function is due to increased or decreased expression of a RIZ or to expression of a mutant RIZ. The test sample can be examined using, for example, immunohistochemical methods as described above or the sample can be further processed and examined. For example, an extract of a test sample can be prepared and examined to determine whether RIZ that is expressed in cells in the sample can associate with Rb in the same manner as RIZ from control cells or whether a variant RIZ, which does not properly associate with Rb, is expressed in the cells in the test sample.

A diagnostic assay kit incorporating a reagent such as an anti-RIZ antibody or Rb can be useful for detecting a pathology due to altered RIZ expression or to expression of a mutant RIZ in a cell. Such a kit is particularly useful because it allows for standardization of assay conditions. A kit can contain, in addition to a reagent, a reaction cocktail that provides suitable reaction conditions for performing the assay and, if desired, a control sample that contains a known amount of RIZ. In addition, the kit can contain an antibody that is specific for the reagent. Where Rb is used as a reagent to detect RIZ, the kit also can contain a competitor molecule such as EIRCEEKPEDL (SEQ ID NO: 6) or EIRCDEKPEDL (SEQ ID NO: 91), which inhibits the

association of RIZ and Rb and, therefore, can confirm the specificity of the binding reaction.

A diagnostic assay should include a simple method for detecting the amount of RIZ in a sample that is bound to the reagent. Detection can be performed by labeling the reagent and detecting the presence of the label using well known methods (see, for example, Harlow and Lane, *supra*, 1988; chap. 9, for labeling an antibody). A reagent can be labeled with various detectable moieties including a radiolabel, an enzyme, biotin or a fluorochrome. Materials for labeling the reagent can be included in the diagnostic kit or can be purchased separately from a commercial source. Following contact of a test sample and, if desired, a control sample, with a labeled reagent, specifically bound reagent can be identified by detecting the particular moiety.

A labeled antibody that can specifically bind the reagent also can be used to identify specific binding of an unlabeled reagent. For example, if the reagent is an anti-RIZ antibody, a second antibody can be used to detect specific binding of the anti-RIZ antibody. A second antibody generally will be specific for the particular class of the first antibody. For example, if an anti-RIZ antibody is of the IgG class, a second antibody will be an anti-IgG antibody. Such second antibodies are readily available from commercial sources. The second antibody can be labeled using a detectable moiety as described above. When a sample is labeled

using a second antibody, the sample is first contacted with a first antibody, then the sample is contacted with the labeled second antibody, which specifically binds to the first antibody and results in a labeled sample.

A method for diagnosing a pathology characterized by an abnormal level of expression of a RIZ can involve measuring the level of expression of a DNA or RNA in the sample. Similarly, diagnosing a pathology characterized by expression of a mutant RIZ or by the presence of a mutant nucleic acid molecule encoding a RIZ can involve detecting the mutation in the RIZ gene or in the RNA encoded by the gene.

For example, a nucleic acid molecule encoding a RIZ can be detected in a test sample using a complementary nucleotide sequence. If desired, the target nucleic acid molecule can be extracted from a test sample by methods well known in the art (See Sambrook et al., *supra*, 1988). Methods to detect the presence of a particular nucleic acid molecule within a population of nucleic acid molecules are well known to those in the art and include, for example, Southern blotting, northern blotting, slot blotting and PCR amplification (see, for example, Sambrook et al., *supra*, 1989). *In situ* hybridization also can be used to identify nucleic acids directly in a sample containing cells or free chromosomes (see, for example, Pardue, in *Nucleic Acid Hybridisation: A practical approach* (IRL Press, 1991), which is incorporated herein by reference).

To detect a nucleic acid molecule encoding a RIZ in a sample, the sample is contacted with the complementary nucleotide sequence, which can hybridize to a nucleic acid molecule encoding the RIZ under relatively stringent conditions. The nucleotide sequence can carry a detectable label such as a radioisotope. The presence of a nucleic acid molecule encoding the RIZ in the sample can be determined, for example, by detecting the level of the specifically bound nucleotide sequence. The normal level of binding of the nucleotide sequence also can be determined in a control sample. An increase or a decrease in the level of nucleic acid molecules encoding a RIZ in the test sample compared to the control sample indicates a pathology characterized by an abnormal expression of the RIZ. A complementary nucleotide sequence for a RIZ can also be used as a primer in a PCR reaction to amplify the RIZ for hybridization by a probe.

A mutant RIZ can be detected by hybridizing with a complementary nucleic acid molecule under relatively stringent conditions essentially as described above except that the complementary sequence is of sufficiently small size to enable selective hybridization to the mutant sequence but not to the wild-type sequence under the conditions chosen for hybridization. Alternatively, the RIZ gene or RNA can be purified directly from a test sample and, if desired, amplified from the sample by PCR and the mutant sequence determined by standard nucleotide sequencing methods (see, for example, Sambrook et al. *supra*, 1989). The mutant nucleic acid encoding a RIZ or the nucleic acid encoding

a mutant RIZ also can be detected in a sample of cells or free chromosomes by *in situ* hybridization techniques (see, for example Pardue, *supra*, 1991).

The following Examples are intended to illustrate but not limit the invention.

EXAMPLE I

Cloning of Mammalian RIZ cDNAs

This section describes methods to clone nucleic acid molecules encoding a RIZ from mammalian cDNA and genomic libraries.

A rat RIZ cDNA was obtained from a rat neonatal cardiac myocyte λgt11 cDNA expression library (Zhu et al., Mol. Cell Biol., 13:4432 (1993), which is incorporated herein by reference). The library was screened using a 56 kD fragment containing the pocket binding site of Rb and the EE epitope (p56EERb) according to previously described methods (Macgregor et al., Oncogene, 5:451-458 (1991), which is incorporated herein by reference).

p56EERb was generated by cloning a synthetic pair of complementary polynucleotides that hybridize to form a double stranded linker encoding the EE-epitope, EEEEYMPME (SEQ ID NO: 8; Grussenmeyer et al., Proc. Natl. Acad. Sci., *USA.*, 82:7952-7954 (1985) and Walter, J. Immune Meth., 88:149-161 (1986), both of which are incorporated herein by reference) and having Bsm I

cohesive ends. The ends of the linker were phosphorylated by T4 kinase and the linker was ligated into the plasmid pET8Rbc (Huang et al., Nature, 350:160-162 (1991), which is incorporated herein by reference) to produce the plasmid p56EERb. The synthetic nucleotides used to make the linker were: 5'-AATCGATGAA GAAGAAGAAT ATATGCCTAT GGAACA-3' (SEQ ID NO: 9), and 5'-TTCCATAGGC ATATATTCTT CTTCTTCATC GATTTG-3' (SEQ ID NO: 10). A clone with four tandem copies of the EE linker was selected and used to direct expression of p56EERb in the *E. coli* strain BL21(DE3)pLysS as previously described (Huang et al., *supra*, 1991).

After induction of 56EERb, the bacterial cells were lysed as described (Huang et al., *supra*, 1991) and 56EERb was precipitated by ammonium sulfate (60% of saturation). The precipitate was dialyzed in dialysis buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF) and subjected to further purification by diethylaminoethyl (DEAE) Sepharose™ chromatography (Pharmacia; Piscataway NJ). Partially purified p56EERb was eluted from DEAE by a salt gradient of 50 mM to 500 mM NaCl. Both the DEAE partially purified fraction and the dialyzed ammonium sulfate precipitate of 56EERb were used for screening the cDNA library.

For binding-specificity control, p56EERb was preincubated with 5 μM poly-L-lysine (Sigma) or 50 μM T- or K-peptide (Huang et al., *supra*, 1991) before applying onto filters. T peptide is an 18 residue synthetic peptide derived from residues 101-118 of SV40 large T

antigen, while K peptide is the same as T peptide except for a lysine residue substituted for a glutamic acid residue at position 107 of SV40 large T antigen (Huang et al., *supra*, 1991). The T peptide binds to the Rb pocket while the K peptide does not.

The binding of p56EERb to a clone expressing a RIZ protein was detected using an anti-EE monoclonal antibody obtained as spent culture medium of the anti-EE hybridoma (Walter, *supra*, 1986) and an alkaline phosphatase conjugated goat anti-mouse IgG antibody specific for mouse immunoglobulin (Promega; Madison WI).

Filters containing 1 x 10^6 library phage plaques were screened using p56EERb and ten positive clones that survived three rounds of plaque purification were selected. Five clones, which maintained their reactivity with p56EERb in the presence of a non-specific inhibitory substance, poly-L-Lysine, but were inhibited from binding p56EERb in the presence of T peptide but not K peptide, were selected for further study. Inhibition by T peptide indicated that the selected clones expressed a product that binds the Rb pocket.

Of the final five clones, four contained an identical 1.9 kilobase (kb) insert. One of the clones, clone 7.1, was subjected to nucleotide sequencing. Sequencing was performed on both strands of the DNA and utilized Sequenase™ (United States Biochemical Corp.; Arlington Hts. IL). Clone 7.1 contained a partial cDNA sequence having a predicted open reading frame encoding

638 amino acids, which formed two types of readily recognizable motifs: a cr2 core motif and 3 zinc finger motifs (see below). The protein encoded by clone 7.1 was designated RIZ for "Rb-interacting zinc finger" protein.

The 1.9 kb insert was used to further screen the cardiac myocyte library and to screen a rat brain B49 cell cDNA library produced in the λZAP vector (Stratagene) according to standard methods (see Sambrook et al., *supra*, 1989) or purchased from a commercial source. Several clones containing overlapping open reading frames were obtained. The overlapping sequences were assembled into a contiguous stretch of 6171 nucleotides to obtain the cDNA sequence for rat RIZ (Figure 1A; SEQ ID NO: 1).

Screening of the rat brain cDNA library also revealed clone 12.1, which contained a nucleotide sequence that encoded an alternative 5'-untranslated sequence as compared to the sequence shown in Figure 1A (see Figure 1B; SEQ ID NO: 94). Thus, at least two forms of RIZ mRNA are expressed in rat brain, suggesting that RIZ mRNA molecules can arise by alternative splicing.

Analysis of the complete rat RIZ cDNA sequence (SEQ ID NO: 1) revealed a large open reading frame beginning at nucleotide 157 and ending at nucleotide 5274 (see SEQ ID NO: 2). The initiation codon at nucleotide 157 was considered the translational start site based on its being the first ATG following an in-frame upstream stop codon at nucleotide 100 and by its match with the

Kozak consensus sequence (Kozak, Nucl. Acids Res. 15: 8125-8148 (1987)). The identity of the start site was confirmed by analyzing an independent cDNA clone that revealed a divergent sequence upstream of nucleotide 92 but otherwise was identical to the assembled cDNA sequence of rat RIZ.

The complete rat RIZ cDNA sequence predicted a protein consisting of 1706 amino acids having a molecular weight of 187,437 Daltons (Figure 1A; SEQ ID NO: 2). Northern blot analysis showed a 7.2 kb major rat RIZ mRNA species. Southern blot analysis indicated that the rat RIZ genome contains a single copy of the RIZ gene. This result, along with the identification of alternative 5'-termini in cDNA molecules encoding rat RIZ, indicates that the different mRNA molecules arise via alternative splicing.

A nucleic acid molecule encoding human RIZ was cloned from a human fetal brain cDNA library (Clonetech; Palo Alto CA) and a human placental genomic cosmid DNA library (Stratagene; San Diego CA) using the rat RIZ cDNA coding regions as a hybridization probe (clone 7.1). The human RIZ cDNA encodes a polypeptide of 1719 amino acids residues (see Figure 9A; SEQ ID NO: 4). In addition to the coding sequence, the human RIZ cDNA contains 5' untranslated as well as a partial 3' untranslated sequence (Figure 9A; SEQ ID NO: 3). Additional 3' untranslated sequence (not determined) for the human RIZ cDNA is located downstream of SEQ ID NO: 3 and upstream of the 3' untranslated sequence shown in Figure 9C (SEQ

ID NO: 97; sequence obtained from Washington University-Merck EST Project; Genbank accession number R56425). The human RIZ gene obtained from the genomic library showed that RIZ coding sequence was divided between eight separate exons.

An allelic variant of the human RIZ gene also was identified. This variant contains a single nucleotide change of T_{969} to A_{969}, leading to a change of amino acid residue D_{283} to E_{283}. The T_{969} allele is estimated to be two times more frequent than the A_{969} allele.

Nucleotide sequences encoding additional human RIZ proteins were isolated from a cDNA library prepared by standard procedures from Y79 retinoblastoma cells. As shown in Figure 9B, cDNA molecules encoding two alternative 5' sequences of human RIZ were identified (SEQ ID NOS: 95 and 96; clones 5Y and 1Y, respectively). Clones 5Y (SEQ ID NO: 95) and 1Y (SEQ ID NO: 96), which replace the first 129 nucleotides shown in Figure 9A (SEQ ID NO: 3), do not encode an ATG initiation codon (see SEQ ID NOS: 102 and 104, respectively). These results indicate that various RIZ proteins can be encoded by alternatively spliced mRNA molecules.

Both the rat and human RIZ proteins have similar sequence motifs including cr1, cr2, ce1, zinc finger, SH3, SH2 and a nuclear localization signal. The

deduced rat and human RIZ amino acid sequences show 83% identity. In addition, a rabbit antiserum prepared to rat RIZ cross reacts with human RIZ.

In addition, a 391 nucleotide sequence encoding a portion of mouse RIZ was obtained by combining sequences of several clones selected from the mouse genomic library 129SVJ (Strategene) by a rat cDNA PR domain probe. The mouse cDNA encodes 130 amino acids (SEQ ID NO: 105) that are equivalent to a.a. positions 77 to 206 of rat RIZ (SEQ ID NO: 2) and a.a. positions 78 to 207 of human RIZ (SEQ ID NO: 4). In particular, the encoded mouse peptide contains blocks B and C of the PR domain peptide. The encoded portion of the mouse PR domain peptide is identical to the corresponding human and rat RIZ PR domain peptide sequences.

EXAMPLE II

Detection and Characterization of RIZ-Rb Binding

This section describes methods for demonstrating binding of RIZ and Rb and for identifying an agent that effectively alters the binding of a RIZ and Rb.

To characterize the interaction between RIZ and Rb, a ^{35}S-labeled fragment of rat RIZ from amino acid position 245-883 (RIZ (245-883)) was produced by subcloning the 1.9 kb insert of clone 7.1 into pBKS+ (Stratagene) to yield plasmid pBKS+7.1. Following subcloning, the RIZ insert was then removed and inserted

downstream of the 5' untranslated sequence of β-globin in the vector pSP64-xβm (Krieg and Melton, Nucl. Acids Res., 12:7057-7070 (1984). SP6 RNA transcripts encoding RIZ (245-883) were produced by linearizing the plasmid encoding this fragment with Sal I and translating the RIZ fragment using a rabbit reticulocyte lysate *in vitro* protein translation system (Promega) containing ^{35}S-methionine. The labeled RIZ fragment had an apparent molecular weight of 125 kD by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which was about 55 kD greater than the predicted molecular mass for this fragment of RIZ. The larger size obtained by SDS-PAGE is likely due to anomalous mobility of the RIZ fragment on the gel.

A 56 kD fragment of Rb produced by bacterial expression from p56Rb plasmid, as described previously (Huang et al., *supra*, 1991), was tested for binding to radiolabeled RIZ (245-883). Binding was detected by immunoprecipitation with an anti-Rb antiserum and Protein A-Sepharose™ (Huang et al., *supra*, 1990), which is incorporated herein by reference; and Huang et al., *supra*, 1991) followed by SDS-PAGE and autoradiography (see Harlow and Lane, *supra*, 1988). The rabbit anti-Rb antiserum was produced to purified p56Rb using previously described methods (see Harlow and Lane, *supra*, 1988).

The amount of binding of RIZ by Rb in the immunoprecipitation reaction was dependent on the

concentration of Rb added. Full binding of ^{35}S-labeled RIZ (245-883) was achieved by 10 nM Rb but not 3.3 nM Rb (not shown). These results indicted that RIZ binds Rb.

A competition experiment was used to compare the relative binding affinity of RIZ for Rb as compared to another Rb binding protein, SV40 large T antigen. The full length large T antigen cDNA was subcloned from Y-62-25-2 into plasmid pSP64 for *in vitro* transcription/translation and ^{35}S-methionine labeling as described above. When approximately equal amounts of T antigen and RIZ were mixed individually or together with the same amount of Rb, similar amounts of T antigen and RIZ, or somewhat more RIZ, were co-precipitated (not shown). These data indicate that RIZ has a similar binding affinity for Rb as does large T antigen.

Several mutations were generated to identify the regions of RIZ that were involved in binding to Rb. A single amino acid substitution was introduced into full length RIZ cDNA in the plasmid pCMVRIZ to change cysteine at a.a. position 307 to glycine. pCMVRIZ was produced by subcloning the full-length RIZ cDNA into the pRc-CMV vector (Invitrogen, San Diego, CA). Mutagenesis of the cr2 motif changing Cys to Gly was performed using the T7 GEN™ mutagenesis kit (United States Biochemical, Arlington Heights, IL) as follows: Briefly, the primer, 5'-CCGGAGATCC GGGCTGAAGA AAAGCCA -3' (SEQ ID NO: 11), was used to direct DNA synthesis on a single stranded antisense template prepared from pBSK-5.4. Vector pBSK-5.4 was produced by cloning the cDNA RIZ amino

terminal clone 5.4 obtained from the B49 λZAP DNA library into vector pBSK+. An Nsi I to Spe I fragment (nucleotide 1-1718) containing the point mutation was sequenced and subcloned into pRc-CMV to produce pCMVmRIZ ($RIZ^{307\text{-}Gly}$). A ^{35}S labeled fragment of RIZ from amino acid position 1-575 (RIZ (1-575)) and ^{35}S-RIZ $(1\text{-}575)^{307\text{-}Gly}$ were produced by *in vitro* transcription/translation of Spe I linearized template as described above.

Binding between labeled RIZ (1-575) and the glycine mutant with 33 nM Rb was evaluated by immunoprecipitation with anti-Rb antiserum followed by SDS-PAGE and autoradiography. The results showed that the 56 kD Rb bound the ^{35}S-RIZ (1-575) fragment but not to the ^{35}S-RIZ $(1\text{-}575)^{307\text{-}Gly}$ cr2 mutant (not shown). These results indicate that the RIZ cr2 motif is involved in Rb binding.

To determine whether the RIZ cr2 motif is functional and sufficient for binding Rb, the 11 amino acid peptide EIRCEEKPEDL (SEQ ID NO: 6), representing a portion of the cr2 motif of RIZ (RIZ-Pep), and a cysteine to glycine mutant of this peptide (RIZ-Pep*) were synthesized according to standard procedures and tested at various concentrations for their ability to inhibit the binding of labeled RIZ (1-575) to 56 kD Rb. Binding was inhibited with wild-type peptide but not the C→G mutant peptide (see Figure 4). These data indicate the cr2 motif of RIZ is sufficient for binding to Rb and that the cysteine at a.a. position 307 in the cr2 motif of RIZ is involved in the binding.

In a similar manner, the binding between radiolabeled RIZ (1-575) and 56 kD Rb was tested for inhibition using the 17 amino acid Rb binding peptide (101-118: T-pep) from the SV40 large T antigen oncoprotein and a position 107 Glu to Lys mutant of T-pep (T-pep*) that lacks Rb binding activity (Huang et al., *supra*, 1990, and Huang et al., *supra*, 1991). Binding was inhibited with T-pep but not with the mutant (Figure 4). These results indicate that RIZ and large T antigen bind to a similar region on Rb.

The 56 kD Rb fragment that binds to RIZ is a C-terminal fragment containing the Rb pocket binding region and a C-terminal extension. To further define the portion of 56 kD Rb that binds to RIZ, several Rb mutant polypeptides were tested for binding to RIZ. Mutant and full length Rb were cloned and *in vitro* transcribed/ translated as described previously (Huang et al., *supra*, 1990). H209 is a point mutation resulting in a single amino acid change in Rb that was identified in the small cell lung cancer H209 cell line (American Type Culture Collection (ATCC) #HTB 172). The various Rb forms were tested for binding to glutathionine S-transferase (GST) fused to a fragment of RIZ from amino acid position 245-573 (GST-RIZ (245-573)). This RIZ fragment contains all of the E1A motifs related to binding Rb and was constructed by cloning a Stu I-Hpa I RIZ fragment (nucleotide 795-3068) into vector pBSK+ to make pBSK+SH. An Eco RI fragment was removed from pBSK+SH and ligated into pGEX-KG to produce vector pKG7.1S containing GST-RIZ (245-573).

The binding between purified GST-RIZ (245-573) and the above radiolabeled Rb wild-type and deletion mutants was determined by immunoprecipitation with an anti-RIZ antiserum followed by SDS-PAGE and autoradiography. The antiserum was generated by injecting rabbits with the purified GST fused to a fragment of RIZ from amino acid position 245-573 (RIZ (245-573)), which contains zinc fingers 1-3, according to commonly used procedures (see Harlow and Lane, *supra*, 1988). GST-RIZ (245-573) used for immunizing rabbits was produced by expression of plasmid pKG7.1S in *E. coli* strain XL-1 blue. The bacteria were lysed and the GST-RIZ fusion protein isolated by glutathionine agarose column chromatography. pKG7.1S was constructed by ligating the 1.9 kb RIZ insert from pB7.1 into vector pGEX-KG. The resulting plasmid was linearized with Spe I, treated with Klenow fragment of DNA polymerase I and religated, thereby introducing a stop codon at the former Spe I site (nucleotide 1876).

The anti-RIZ antiserum specifically bound to *in vitro* translated RIZ (245-883) expressed from pB7.1. This binding was inhibited by the addition of the immunogen, GST-RIZ (245-573).

Purified GST-RIZ (245-573) showed binding to wild-type Rb and the B3 mutant of Rb, which contains a deletion C-terminal to the Rb binding pocket, but failed to bind three different forms of Rb having a deletion within the pocket (Figure 5A). These results indicate that the Rb pocket, which was initially defined for its

role in binding of oncoproteins such as the large T antigen or E1A, also is required for binding to RIZ. RIZ-binding by Rb does not require the C-terminal sequence distal to the pocket, as do certain cellular proteins such as E2F (see Huang et al., *DNA Cell Biol.*, 11:539-548 (1992); Qin et al., Genes Devel., 6:953-964 (1992)) and c-Abl oncoprotein (see Welch and Wang, Cell 75:779-790 (1993)). The binding results map the C-terminal boundary of the RIZ-binding domain of Rb to residue 803 of Rb, in close proximity to the beginning of the C-terminal boundary of the Rb pocket (Figure 5B).

Rat RIZ was tested for binding to Rb in HT1080 cells (ATCC #ICCL 121). The cells were grown in DMEM supplemented with 10% fetal calf serum. Cells were lysed in ELB buffer (50 mM HEPES, pH 7.5, 250 mM NaCl, 0.1% NP-40) supplemented with 5 mM EDTA, 50 mM NaF, 1 mM Na orthovanadate, 1 mM of DTT, aprotinin, leupeptin, and PMSF. The lysate was cleared of cell debris by centrifugation at 12,000 rpm for 10 min in a microfuge.

Binding between 4 μg GST-RIZ (215-462) and Rb from HT1080 cell extract was evaluated by first mixing the two, then binding GST-RIZ and any associated Rb to glutathione-agarose beads. The beads were washed in binding buffer and the bound complexes were eluted by boiling in SDS buffer and analyzed by immunoblotting with anti-Rb antiserum. Immunoblotting was performed by standard techniques (see, for example, Harlow and Lane, *supra*, 1988).

GST-RIZ (215-462) bound to the fastest migrating forms of Rb, representing hypophosphorylated Rb (Figure 6, lane 2). The specificity of the interaction between RIZ and hypophosphorylated Rb was demonstrated by showing that the addition of a source of E1A protein inhibited binding (Figure 6, lane 3). A cell lysate from 293 stably transfected to express E1A was used as the source of E1A.

EXAMPLE III

Structural and Functional Comparison Between RIZ and E1A

The similarity in sequence of particular domains between RIZ and Adenovirus E1A (see Figure 2A) and the shared property of Rb binding indicated significant structural similarity between RIZ and E1A. To investigate this relationship further, the anti-RIZ antiserum raised against the GST-RIZ (245-573) fusion protein containing the cr2, ce1 and part of the cr1 motifs, was tested for cross reactivity with E1A. For these experiments, E1A was labeled with ^{35}S-methionine during *in vitro* transcription/translation using methods described above.

Anti-RIZ antiserum cross reacted weakly with E1A (not shown). To further verify binding between anti-RIZ and E1A, the cross reactive antibodies from the anti-RIZ antiserum were purified by affinity chromatography on a column containing E1A 12S protein. The column was prepared by coupling Affi-gel 10™ beads (Bio-Rad Labs; Hercules CA) with the purified GST-E1A 12S fusion protein

expressed from pGSTE1A12S (Taylor et al., Mol. Cell. Biol. 13:4714-4727 (1993), which is incorporated herein by reference). Antibody affinity purification was conducted by high pH elution according to standard procedures (see Harlow and Lane, *supra*, 1988).

Anti-RIZ antibodies purified from the E1A affinity column were tested for binding to RIZ and E1A. Both proteins were bound by the antibodies, confirming the original cross reactivity of the anti-RIZ antiserum with E1A 12S (not shown). The E1A-affinity purified RIZ antibodies were designated "anti-ce1" for cross reacting E1A antigen.

Anti-ce1 antibodies were tested for binding to various deletion mutants of RIZ and E1A 12S in order to map the location of the ce1 epitope on each molecule. A RIZ mutant truncated after residue 304 (RIZ304) was produced by *in vitro* transcription/translation of a BAM HI digested fragment derived from a BAM H1 mutant of pCMVRIZ. A T7 GEN™ mutagenesis kit (U.S. Biochemical) was used to introduce a Bam HI restriction site into pCMVRIZ at RIZ nucleotide 1067 using the primer 5'-TTCACACCGG ATCCCCGGCT CTTTCGC -3' (SEQ ID NO: 12). The Bam HI fragment was then excised and cloned into pRc-CMV to yield a vector encoding RIZ304.

A RIZ mutant truncated after residue 318 (RIZ318) was produced by PCR using full-length RIZ as the template and an upstream T7 primer (Stratagene) and a downstream RIZ primer 5'- TGGCTCTTCT AATAAGTC -3' (SEQ ID

NO: 13). The PCR fragment was cloned into pCRSK+ (Stratagene) and used to produce the RIZ318 polypeptide by *in vitro* T7 transcription/translation.

E1A 12S, truncated at residue 223 (E1A223) was produced by generating a PCR fragment of E1A using an upstream SP6 primer (Stratagene) a downstream E1A primer 5'- GATACATTCC ACAGCCTG -3' (SEQ ID NO: 19) and the plasmid pGEM1Ad5E1A12S as template. The resulting PCR fragment was cloned into pCRSK+, which was used to direct the synthesis of the mutant E1A 12S protein by SP6 *in vitro* transcription/translation. The full length E1A 12S (E1A243) was produced from vector pGEM1Ad5E1A12S by *in vitro* transcription/translation as described above for the other vectors.

Anti-ce1 antibody bound to RIZ truncated at residue 318 but failed to react with RIZ truncated at residue 304 (not shown). These results indicate that the ce1 cross reactive antigenic determinant lies within residues 304 to 318 of RIZ. Anti-ce1 antibody bound to full length E1A (EIA243) but failed to react with the C-terminal deletion mutant of E1A (E1A223; not shown). These results indicate that the ce1 epitope is located within the C-terminal 20 amino acids of E1A 12S.

The regions of RIZ and E1A 12S that contain the ce1 epitope show significant amino acid sequence homology (Figure 2A). The sequence ^{312}EDLLEE (SEQ ID NO: 20) in RIZ and the sequence ^{224}EDLLNE (SEQ ID NO: 21) in E1A are likely sites for the ce1 epitope. To evaluate this

possibility, an 11 amino acid peptide encompassing residues 310-320 in RIZ (ce1 peptide) KPEDLLEEPQS (SEQ ID NO: 7) and an overlapping 11 amino acid control peptide encompassing residues 304-314 containing the cr2 core motif of RIZ, peptide EIRCEEKPEDL (SEQ ID NO: 6), were synthesized by solid phase peptide synthesis and tested for their ability to block binding between anti-ce1 antibody and RIZ or E1A.

The ce1 peptide inhibited binding between anti-ce1 antibody and either ^{35}S-RIZ318 or ^{35}S-E1A 12S (E1A243); the cr2 peptide was not inhibitory (not shown). These experiments indicated that the ce1 epitope is located in the sequence ^{312}EDLLEE (SEQ ID NO: 20) in RIZ and the homologous sequence ^{224}EDLLNE (SEQ ID NO: 21) in E1A.

Anti-ce1 was tested for binding to a preformed RIZ-Rb complex to determine if the ce1 epitope is directly involved or closely associated with regions in the RIZ-Rb binding interface. In these experiments, ^{35}S-labeled full-length Rb was preincubated with *in vitro* translated RIZ (215-462) to form the RIZ-Rb complex prior to adding anti-ce1 antibody for immunoprecipitation. In these experiments, the GST portion of GST-RIZ (215-462) had been previously removed by thrombin cleavage and was purified from any residual uncleaved fusion protein by adsorption with glutathionine-agarose.

The anti-ce1 antibody bound to the preformed RIZ-Rb complex (not shown). Although the binding could

be characterized as weak, this was similar in reactivity with anti-ce1 binding with RIZ. Because no evidence of RIZ homo-oligomer formation was observed, Rb likely interacts directly with RIZ that also was bound by anti-ce1. The failure to observe homo-oligomer formation was based on the lack of binding between GST-RIZ (215-462) and ^{35}S-labeled full length RIZ.

The above binding study also was performed in reverse order by first precomplexing ^{35}S-labeled RIZ (1-575) with full-length Rb, then testing the complex for binding to anti-ce1 antibody. The result showed that the RIZ fragment bound anti-ce1 antibody regardless of whether RIZ had complexed with Rb (not shown). These experiments indicate that the ce1 epitope is not significantly involved in the interface between RIZ and Rb in the RIZ-Rb complex.

EXAMPLE IV

DNA- And GTP-Binding Activities of RIZ

To evaluate whether the zinc finger domains of RIZ can bind to DNA, the RIZ finger motifs 1 to 3 from amino acid position 245-573 or finger 4 to 6 from amino acid position 1114-1260 were expressed as GST fusion proteins, GSTZ13 and GSTZ46, respectively. The GST-RIZ fragments were purified by glutathionine agarose chromatography (Guan and Dixon, Anal. Biochem. 192:262-267 (1991), which is incorporated herein by reference) and evaluated for binding in a filter-based DNA-binding

assay (Sukegawa and Blobel, Cell 72:29-38 (1993), which is incorporated herein by reference). To obtain GSTZ46, a fragment encoding RIZ (1114-1260) was made by PCR using primers 5'- GTGGTCCAAG AAACATTC -3' (SEQ ID NO: 17) and 5'- TCGTGTAAAG CTCTTCAG -3' (SEQ ID NO: 18) and pCMVRIZ as template. The PCR fragment was cloned into pBKS+, then into pGEX-KG (Guan and Dixon, *supra*, 1991).

The filter-based DNA binding assay was performed by electrophoresing 0.5 μg of purified GST or GST-RIZ fusion proteins by SDS-PAGE and transferring the proteins to nitrocellulose. The proteins were renatured by incubating the nitrocellulose for 3 hr in binding buffer (50 mM Tris-HCl, pH 8, 100 mM KCl, 0.1% Triton X-100™, 10% glycerol, and 0.1 mM $ZnCl_2$). ^{32}P-labeled, randomly sheared rat ovary genomic DNA was added to the buffer and the nitrocellulose was incubated for an additional 3 hr. Blots were washed 5 times in binding buffer, dried, then autoradiographed. In some experiments, the binding buffer contained 10 mM EDTA and 2 mM DTT but no $ZnCl_2$.

The DNA filter binding assay showed that GSTZ13, containing zinc fingers 1 to 3 bound to rat DNA while GSTZ46, containing zinc fingers 4 to 6 did not bind (Figure 7A and 7B). In addition, RIZ fragment containing zinc fingers 1-3 bound to DNA in a Zn^{++} ion dependent manner (Figure 7). These results indicate that RIZ zinc finger domains 1 to 3 are active in binding DNA.

The GTPase domain of RIZ, which was defined by sequence homology, was evaluated to determine if it was functionally active. For these studies, a fragment of RIZ from amino acid position 760-949 (RIZ 760-949), containing the putative GTPase domain was expressed as a fusion to GST from the plasmid pKG-G and tested for binding to radiolabeled nucleotides. pKG-G was produced by PCR amplification of the nucleotide sequence encoding RIZ (760-949) using primers 5'- TCTCCACAGC ACAGCCCT -3' (SEQ ID NO: 15), and 5'- GGATAAGGAG GCTGTCTG -3' (SEQ ID NO: 16) and pCMVRIZ as template. The fragment was cloned into pBSK+ and then into pGEX-KG, expressed and purified by glutathionine-agarose as described above. GST was also expressed from vector pGEX-KG and purified as described above.

To measure GTP-binding, 0.5 μg of GST-RIZ or control GST proteins were separated by SDS-PAGE and blotted onto nitrocellulose. Proteins were renatured in GTP-binding buffer (50 mM Tris-HCl, pH8, 100 mM KCl, 10% glycerol, 0.1% Triton X-100, and 2 mM $ZnSO_2$). The nitrocellulose was incubated for 30 min in GTP-binding buffer and then for 2 hr in GTP-binding buffer containing 1 μM α-^{32}P-GTP (800 Ci/mmol). The nitrocellulose was washed 5 times in GTP-binding buffer, dried and autoradiographed. In some samples, 20 mM unlabeled nucleotides were incubated with the nitrocellulose for 1 hr prior to the addition of α-^{32}P-GTP.

The RIZ GTPase fusion protein (GST-G), but not the control GST protein, bound to radiolabeled GTP (Figure 8A and lanes 1 and 2 of Figure 8B). Binding was specific for GTP, as an excess amount of unlabeled GTP inhibited binding of RIZ GTPase to radiolabeled GTP but excess unlabeled ATP, CTP, or UTP did not effect binding (Figure 8B, lanes 3-6). These data indicate that the GTPase domain of RIZ is functionally active.

EXAMPLE V

Expression of RIZ in Cells, Tissues and Organs

This example provides methods to identify nucleic acid molecules encoding a RIZ in mammalian cells, tissues and organs.

RNA samples were obtained from rat tissues and from the mouse pituitary cell line Att-20 (ATCC #CCL 89) by extraction with RNAzol (Biotecx; Houston TX) following manufacturer's procedures and purification of the mRNA by oligo dT cellulose chromatography using an oligo dT mRNA kit (Qiagen) using standard procedures. mRNA was also extracted as described above from a variety of human cell lines obtained from the American Type Culture Collection (Rockville MD). Northern blotting was performed using these various mRNAs and hybridization with a ^{32}P-labeled rat RIZ (representing a.a. positions 245-883) according to standard procedures (Sambrook et. al., *supra*, 1989).

Northern blotting showed a major 7.2 kb major RIZ mRNA species primarily localized to rat neuroendocrine tissues (Figure 11A). The testes showed a 5 kb mRNA species, which is smaller than the RIZ mRNA detected in the other organs or tissues.

Further evaluation of mRNA expression was performed using an RNase protection method. The method was performed according to standard procedures using a ^{32}P-labeled rat RIZ (representing a.a. position 463-574) as the probe. The results showed abundant levels of RIZ mRNA in various neural tissues of a 16 day rat fetus as well as the placenta (Figure 11B). In contrast, little if any mRNA was detected in adult rat tissues by this method.

RNase protection showed that RIZ mRNA was detectable in the human retinoblastoma cell lines, Y79 and Weri-Rb-1, with lower levels of detection in a variety of other human cell lines (Figure 12). These results indicate that RIZ mRNA is expressed in large amounts in neuroendocrine related tissues of mammals and can be involved in fetal development.

Several segments of human RIZ cDNA, encompassing the full length coding region were used as probes to screen a human placental genomic cosmid library to isolate the RIZ gene. Several genomic clones were isolated and the segments encoding RIZ were localized within the clones by restriction mapping and nucleotide

sequencing. The genomic clones showed that the sequence encoding RIZ is distributed across eight exons in the gene, with the majority of RIZ sequence (4.3 kb) contained in exon 7.

EXAMPLE VI

Analysis of the RIZ Gene in Normal and Tumor Cells

This example provides methods to detect the RIZ gene by direct chromosomal analysis and to evaluate mutations in the RIZ gene in tumor cells.

A. Chromosomal Localization of the Human RIZ Gene

A cosmid clone with a 35 kb insert that contains exons 7 and 8 was used as a probe for fluorescence *in situ* hybridization (FISH) on R-banded metaphase chromosomes to detect the chromosomal localization of the human RIZ gene. The method for FISH was performed as described previously (Takahashi et al., Hum. Genet. 88:119-121 (1991), which is incorporated herein by reference). Cot-1 DNA (BRL; Gaithersburg MD) was used for the suppression of repetitive sequences present in this clone according to methods described by Lichter et al., (Lichter et. al., Proc. Natl. Acad. Sci., *USA* 87:6634-6638 (1990), which is incorporated herein by reference) using a 20 fold excess of Cot-1 DNA. Ektachrome film (Kodak, ASA100) was used for the microphotography (filter combination, Nikon B-2A).

Of 100 R-banded metaphase plates evaluated by the FISH method, 52 plates showed hybridization of the probe to both chromatids of chromosome 1 at band p36.13-p36.23, 44 plates showed hybridization of the probe only to one chromatid of chromosome 1, and four plates showed no hybridization.

Further localization of the RIZ gene to chromosome 1p36 was accomplished at the molecular level by YAC cloning. A CEPH-derived human mega-YAC library (CEPH, France) was screened by PCR using two oligonucleotide primers to amplify a 290 bp fragment within the RIZ exon 7. YAC DNA was amplified in a total volume of 10 μl containing 1xPCR buffer (50mM KCl/10 mM Tris-HCl, pH 8.3/1.5 mM $MgCl_2$), 200 μM of each dNTP, 0.3 μM of each primer (SSO 81: 5'CCAGAACCAGACGAGCGATT3' (SEQ ID NO: 92) and SSO 82: 5'AGTTCTGGGGATTTGCATG3' (SEQ ID NO: 93)), 0.2 U Taq DNA polymerase (Perkin Elmer, Norwalk, CT). One of the primers was end-labeled using ^{32}P-γ-ATP and T4 polynucleotide kinase. The PCR fragments were analyzed by acrylamide gel electrophoresis followed by autoradiography.

PCR screening of the CEPH human mega-YAC library for RIZ gene sequences identified two YAC clones, 796H4 and 807H7. A search of Genethon human genome database (Genethon; Paris France) showed that the clones contained the polymorphic marker D1S228, which maps to chromosome 1p36 (Gyapay et al., Nat. Genet. 7:246-339 (1994); Weissenbach et al., Nature 359:794-801 (1992) and

Cohen et al., Nature 366:698-701 (1993), each of which is incorporated herein by reference)). Thus, these results indicate that the RIZ gene is localized to chromosome 1p36.

B. Allele-Specific RIZ Expression in Human Melanoma Cells

Genomic DNA from blood and placental samples of normal subjects and tumor cell lines were prepared by incubating cells for 1 hr at 55 °C in 50 mM Tris-HCl, pH 8.0/100 mM EDTA/0.5% SDS/500 μg/ml Proteinase K. After phenol/chloroform and chloroform extraction, the DNA was precipitated. RNA was prepared from a pellet of 5-10 x 10^6 PBS washed cells or from fetal tissues using RNAzol (Biotecx Laboratories; Houston TX).

Southern blots were performed on melanoma cell DNA using a 1 kb human RIZ cDNA probe (1.1). The results showed identification of the appropriate sized bands for the RIZ gene in the melanoma cells, indicating no gross abnormalities in the RIZ gene in these cells.

To determine whether both alleles of RIZ were active in melanoma cells, the frequencies of the two allelic variants of RIZ were determined for melanoma and compared with the frequency in the population. RIZ genotyping was performed by amplification of a 290 bp fragment representing RIZ (a.a. residues from about 230-330) using PCR on genomic DNA isolated from 28 normal individuals and 69 human melanoma cell lines. PCR amplification of the 290 bp fragment was performed on 100

ng of genomic DNA in a total volume of 50 μl containing 1x PCR buffer, 200 μM of each dNTP, 0.3 μM of each primer (SSO 81 + SSO 82), and 1 U of Taq DNA polymerase (Perkin Elmer). The PCR product was sequenced to determine the codons encoding RIZ a.a. position 283. The RIZ D283 allele encodes an Asp residue at a.a. position 283 by the codon GAT, while the RIZ E283 allele encodes a Glu residue at a.a. position 283 by the codon GAA.

Genotyping of DNA from 28 normal individuals showed that fifteen were homozygous for the RIZ D283 allele (53%) three were homozygous for the E283 allele (10%) and ten were heterozygous (35%). Thus, the overall frequency of the RIZ E283 allele in the population of normal individuals studies was estimated to be about 28.5%.

Genotyping the DNA of 69 melanoma cell lines showed that 40 were homozygous for the RIZ D283 allele (58%), nine were homozygous for RIZ E283 (13%) and 20 were heterozygous (29%).

RNA samples from the 20 heterozygous melanoma cell lines were sequenced to determine if both alleles were transcribed in the cell or if only one allele was transcribed (ie. monoallelic expression). Sequencing was performed on DNA products produced by reverse transcription-PCR (RT-PCR) amplification using specific RIZ primers. RT-PCR amplification was performed according to the manufacturer's instructions (GeneAmp RNA PCR kit; Perkin Elmer). A 640 bp fragment encoding RIZ

exons 5-7 was obtained from transcription of 1 μg of total RNA using the SSO 82 primer (SEQ ID NO: 93) and PCR amplification using the SSO 24 primer (5'GCGAGGAGCTCCTGGTCTGG3'; SEQ ID NO: 106) and the SSO 82 primer (SEQ ID NO: 93). The amplified fragment was gel purified and sequenced using primer SSO 82 and a CircumVent™ Thermal Cycle DNA Sequencing kit (New England Biolabs; Beverly MA). The sequencing products were analyzed on a 6% sequencing gel.

Sequencing of amplified and transcribed RIZ mRNA from 20 heterozygous melanoma cell lines showed that transcripts representing one of the RIZ alleles were reduced or absent in four of the cell lines. These cell lines were D283/E283 RIZ heterozygotes (ie. designated 5-5/6-4). The D283 allele transcript was not detected in SK-MEL-14 cells (ATCC) while the E283 RIZ transcript was not detected in WM983C and WM1361C cell lines. The SK-MEL-23 cell line expressed reduced levels of the D283 RIZ allele. In contrast to the melanoma results, sequencing of amplified and transcribed mRNA from RIZ heterozygotes representing seven non-melanoma cell lines and two normal human placental tissues showed no loss or reduction of RIZ alelic expression.

The 4 melanoma cell lines heterozygous for RIZ were evaluated to determine the amount of RIZ protein produced by the cells. RIZ protein level was estimated qualitatively by immunoprecipitation of RIZ from cell extracts with anti-RIZ antibody followed by immunoblotting the isolated RIZ with the anti-RIZ

antibody. The melanoma cell line SK-MEL-23 produced about 50% less RIZ protein than the other melanoma cell lines tested. These data indicate that the reduction in expression of the RIZ D283 allele in SK-MEL-23 resulted in a decrease in overall expression of RIZ protein in the cell.

The SK-MEL-23 melanoma cell line was cultured for 3 to 6 days with 3 uM of the demethylating agent 5-azacytidine. The reduced expression of the RIZ D283 transcripts in these cells was unaffected, indicating that the reduced levels of RIZ transcription were not due to increased DNA methylation.

SK-MEL-23 melanoma cells were transfected with an expression vector encoding full length RIZ cDNA (pCMVRIZ)to determine if increasing the level of RIZ expression can reduce the growth potential of the cells. RIZ transfected SK-MEL-23 cells showed increased expression of RIZ in the cell nucleus by immunostaining with monoclonal antibody D27 and, showed a reduced ability to form colonies *in vitro*. These results indicate increasing the level of RIZ in tumor cells that are deficient in RIZ expression can reduce the growth of the tumor cells.

EXAMPLE VII

RIZ Protein Represses Transcription

This example demonstrates that RIZ or an active fragment of a RIZ can repress transcription.

A fusion protein containing RIZ or deletion mutants of RIZ fused to the DNA binding domain of GAL4 was produced by subcloning the rat RIZ cDNA (SEQ ID NO: 1) or portions thereof into the plasmid pSG424 (Sadowski and Ptashne, Nucl. Acids Res. 17:7359 (1989), which is incorporated herein by reference). The plasmid encoding the RIZ/GAL4 fusion protein was transfected into CV1 cells, COS cells or C33A cells (ATCC) along with a plasmid containing a CAT reporter gene linked to the thymidine kinase promoter and GAL4 promotor (Shi et al., Cell 67:377-388 (1991), which is incorporated herein by reference).

Reporter gene (CAT) expression was decreased in cells transfected with the RIZ/GAL4 (containing RIZ a.a. 17-1706; SEQ ID NO: 2) as compared to the level of CAT expression in cells transfected with plasmid pSG424, but lacking the RIZ sequence. The repressor activity of RIZ was mapped to amino acids 17 to 900 from the amino terminus (SEQ ID NO: 2). The PR domain is required for transcriptional repression but, alone, is not sufficient for full repressor activity; a region between a.a. positions 573 to 900 also is required. These results demonstrate that RIZ or an active fragment thereof can act as a transcriptional regulator. Furthermore, the transcriptional repressor role of RIZ is independent of the regulatory effect RIZ has due to its interaction with Rb protein because C33A cells that were repressed by RIZ/GAL4 do not express Rb.

CAT reporter activity was decreased by transfection of RIZ/GAL4 containing amino acids 71-1706 (SEQ ID NO:2), however the amount of repressor activity was less than the RIZ/GAL4 containing RIZ amino acids 17-1706 (SEQ ID NO: 2). This result indicates that the PR domain is involved in RIZ-mediated transcriptional regulation and that absense of block A of the PR domain reduces but does not eliminate the RIZ repressor activity.

Although the invention has been described with reference to the above-provided examples, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the claims that follow.

I claim:

1. A substantially purified nucleic acid molecule encoding a mammalian RIZ.

2. A vector, comprising the nucleic acid molecule of claim 1.

3. A host cell, containing the vector of claim 2.

4. The nucleic acid molecule of claim 1, wherein said RIZ is human RIZ having substantially the nucleotide sequence shown in Figure 9A (SEQ ID NO: 3).

5. The nucleic acid molecule of claim 4, wherein an adenine replaces a thymine at nucleotide position 969.

6. The nucleic acid molecule of claim 1, wherein said RIZ is human RIZ having substantially the nucleotide sequence shown in Figure 9B (SEQ ID NO: 95) and the nucleotide sequence from nucleotide position 130 to 5277 shown in Figure 9A (SEQ ID NO: 3).

7. The nucleic acid molecule of claim 1, wherein said RIZ is human RIZ having substantially the nucleotide sequence shown in Figure 9B (SEQ ID NO: 96) and the nucleotide sequence from nucleotide postion 130 to 5277 shown in Figure 9A (SEQ ID NO: 3).

8. A nucleic acid molecule encoding a human RIZ having substantially the amino acid sequence shown in Figure 9A (SEQ ID NO:4).

9. A nucleotide sequence, comprising at least ten nucleotides that specifically hybridizes under relatively stringent conditions to the nucleic acid molecule of claim 4 or to a nucleic acid molecule complementary to said nucleic acid molecule of claim 4.

10. The nucleic acid molecule of claim 1, wherein said RIZ is rat RIZ having substantially the nucleotide sequence shown in Figure 1A (SEQ ID NO: 1).

11. A nucleic acid encoding a rat RIZ having substantially the amino acid sequence shown in Figure 1A (SEQ ID NO: 2).

12. A nucleotide sequence, comprising at least ten nucleotides that specifically hybridizes under relatively stringent conditions to the nucleic acid molecule of claim 11 or a nucleic acid molecule complementary to said nucleic acid molecule of claim 11.

13. A nucleic acid molecule encoding a mouse RIZ having substantially the amino acid sequence shown in Figure 14 (SEQ ID NO: 105).

14. A substantially purified mutant nucleic acid molecule encoding a RIZ.

15. A nucleotide sequence, comprising at least ten nucleotides that specifically hybridizes under relatively stringent conditions to the mutant nucleic acid molecule of claim 14 or to a nucleic acid molecule complementary to said mutant nucleic acid molecule of claim 14 but not to the nucleic acid molecule of claim 1.

16. A substantially purified mammalian retinoblastoma protein interacting zinc finger protein (RIZ).

17. A substantially purified active fragment of a mammalian RIZ.

18. The active fragment of claim 17, comprising substantially the amino acid sequence EIRCEEKPEDL (SEQ ID NO: 6).

19. The active fragment of claim 17, comprising the amino acid sequence EIRCDEKPEDL (SEQ ID NO: 91).

20. Substantially purified human RIZ, having substantially the amino acid sequence shown in Figure 9A, (SEQ ID NO: 4).

21. The human RIZ of claim 20, wherein glutamic acid is substituted for aspartic acid at amino acid position 283.

22. Substantially purified rat RIZ, having substantially the amino acid sequence shown in Figure 1A, (SEQ ID NO: 2).

23. An antibody that specifically binds to the RIZ of claim 16.

24. A substantially purified mutant human RIZ or active fragment thereof.

25. An antibody that specifically binds to the mutant RIZ of claim 24.

26. A PR domain peptide, comprising the amino acid sequence: Y-A-X100-B-X101-C-Z,

wherein Y is about 8 to 13 independently selected amino acids;

X100 is about 20 to 35 independently selected amino acids;

X101 is about 20 to 35 independently selected amino acids;

Z is about 8 independently selected amino acids;

A is Ile-X2-X3-Gly-X4-X5-X6-Gly-X7-X8-X9-X10,

wherein X6 is Phe or Ile, X7 is Pro or Val, X10 is Gly or Lys, and X2, X3, X4, X5, X8 and X9 each is one independently selected amino acid;

B is Ile-X11-X12-X13-X14-X15-X16-X17-X18-X19-X20-X21,

wherein X11 is Asp or Cys, X15 is Pro or Glu, X16 is Glu or Asp, X20 is Trp or Phe, X21 is Met or Leu, and X12, X13, X14, X17, X18 and X19 each is an independently selected amino acid; and

C is X22-X23-L-X24-X25-X26-X27-X28-X29-X30-D, wherein X22 is Glu or Gln, X23 is Glu or Asn, X24 is Leu or Ile, X25 is Val or Leu, X30 is Arg or Glu, X27 and X29 each is an independently selected amino acid, and X26 and X28 each independently is absent or is an independently selected amino acid.

27. The PR domain peptide of claim 26, wherein

in A, X3 is Lys or Val, X5 is Arg or Lys, X6 is Phe, X7 is Pro, X9 is Ile or Val, and X10 is Gly;

in B, X11 is Asp, X12 is Gly or Ala, X17 is Lys or Val, X18 is Ser or Gly, X19 is Asn or Ser, and X20 is Trp; and

in C, X23 is Glu, X24 is Leu, X26 is Trp or Phe, X27 is Tyr or Met, and X28, when present, is Asn or Lys.

28. The PR domain peptide of claim 26, wherein said amino acid sequence is SEQ ID NO: 99.

29. The PR domain peptide of claim 26, wherein said amino acid sequence is shown as amino acid positions 36 to 151 in SEQ ID NO: 2.

30. The PR domain peptide of claim 26, wherein said amino acid sequence is selected from the group consisting of SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 100 and SEQ ID NO: 101.

31. The PR domain peptide of claim 26, wherein A is selected from the group consisting of amino acid positions 71 to 81 of PRDI-BF1 (SEQ ID NO: 98), amino acid positions 50 to 61 of RIZ (SEQ ID NO: 99), amino acid positions 71 to 82 of EVI-1 (SEQ ID NO: 100) and amino acid positions 29 to 40 of egl-43 (SEQ ID NO: 101).

32. The PR domain peptide of claim 26, wherein B is selected from the group consisting of amino acid positions 112 to 123 of PRDI-BF1(SEQ ID NO: 98), amino acid positions 88 to 99 of RIZ (SEQ ID NO: 99), amino acid positions 108 to 119 of EVI-1 (SEQ ID NO: 100) and amino acid positions 65 to 76 of egl-43 (SEQ ID NO: 101).

33. The PR domain peptide of claim 26, wherein C is selected from the group consisting of amino acid positions 158 to 167 of PRDI-BF1 (SEQ ID NO: 98), amino acid positions 134 to 44 of RIZ (SEQ ID NO: 99), amino acid positions 154 to 164 of EVI-1 (SEQ ID NO: 100) and amino acid positions 107 to 116 of egl-43 (SEQ ID NO: 101).

34. A PR domain peptide, comprising an amino acid sequence selected from the group consisting of:

Ile-X2-X3-Gly-X4-X5-X6-Gly-X7-X8-X9-X10,

wherein X6 is Phe or Ile, X7 is Pro or Val, X10 is Gly or Lys, and X2, X3, X4, X5, X8 and X9 each is one independently selected amino acid;

Ile-X11-X12-X13-X14-X15-X16-X17-X18-X19-X20-X21,

wherein X11 is Asp or Cys, X15 is Pro or Glu, X16 is Glu or Asp, X20 is Trp or Phe, X21 is Met or Leu, and X12, X13, X14, X17, X18 and X19 each is an independently selected amino acid; and

X22-X23-L-X24-X25-X26-X27-X28-X29-X30-D,

wherein X22 is Glu or Gln, X23 is Glu or Asn, X24 is Leu or Ile, X25 is Val or Leu, X30 is Arg or Glu, X27 and X29 each is an independently selected amino acid, and X26 and X28 each independently is absent or is an independently selected amino acid.

35. The PR domain peptide of claim 34, wherein said amino acid sequence is selected from the group consisting of amino acid positions 50 to 61 of RIZ (SEQ ID NO: 99), amino acid positions 88 to 99 of RIZ (SEQ ID NO: 99) and amino acid positions 134 to 44 of RIZ (SEQ ID NO: 99).

36. The PR domain peptide of claim 34, wherein said amino acid sequence is selected from the group consisting of amino acid positions 71 to 81 of PRDI-BF1 (SEQ ID NO: 98), amino acid positions 71 to 82 of EVI-1 (SEQ ID NO: 100), amino acid positions 29 to 40 of egl-43 (SEQ ID NO: 101), amino acid positions 112 to 123 of PRDI-BF1(SEQ ID NO: 98), amino acid positions 108 to 119 of EVI-1 (SEQ ID NO: 100), amino acid positions 65 to 76 of egl-43 (SEQ ID NO: 101), amino acid positions 158 to 167 of PRDI-BF1 (SEQ ID NO: 98), amino acid positions 154 to 164 of EVI-1 (SEQ ID NO: 100) and amino acid positions 107 to 116 of egl-43 (SEQ ID NO: 101).

37. A method of identifying an effective agent, which alters the association of a RIZ with a second molecule, comprising the steps of:

a. contacting the RIZ and the second molecule under suitable conditions, which allow said RIZ and said second molecule to bind, with an agent suspected of being able to alter the association of said RIZ with said second molecule; and

b. detecting the altered association of said RIZ with said second molecule, wherein said altered association identifies an effective agent.

38. The method of claim 37, wherein contacting step a. comprises:

a1. contacting the RIZ and the second molecule under suitable conditions, which allow said RIZ and said second molecule to bind;

a2. thereafter contacting said RIZ and said second molecule with an agent suspected of being able to alter the association of said RIZ with said second molecule.

39. The method of claim 37, wherein said second molecule is Rb protein or a fragment of Rb protein that can associate with a RIZ.

40. The method of claim 37, wherein said second molecule is a nucleic acid.

41. The method of claim 37, wherein said effective agent is a peptide.

42. The method of claim 37, wherein said effective agent is a mutant RIZ.

43. The method of claim 37, wherein said RIZ and said second molecule are in a test sample.

44. A method for altering the association of a RIZ with a second molecule in a cell, comprising contacting the cell with an effective agent identified using the method of claim 37.

45. A method for modulating a function of a cell, comprising contacting the cell with an effective agent identified using the method of claim 37.

46. The method of claim 45, wherein said function is cell proliferation.

47. A method for modulating a function of a cell, comprising the steps of:

a. introducing a nucleic acid molecule encoding a RIZ into said cell; and

b. expressing said RIZ in said cell, wherein said RIZ modulates a function of said cell.

48. A method for promoting the growth of a cell, comprising the steps of:

a. introducing a nucleic acid molecule encoding an active fragment of a RIZ into the cell, wherein said fragment can bind Rb but lacks the growth suppressing properties of a RIZ; and

b. expressing said active fragment in said cell, wherein said active fragment promotes the growth of said cell.

49. The method of claim 48, wherein said cell is in a subject.

50. The method of claim 48, wherein said active fragment is selected from the group, consisting of EIRCEEKPEDL (SEQ ID NO: 6) and EIRCDEKPEDL (SEQ ID NO: 91).

51. A method for reducing the growth of a tumor cell having a mutant RIZ in a subject, comprising the steps of:

a. introducing a nucleic acid molecule encoding a RIZ into the tumor cell; and

b. expressing said RIZ in said tumor cell, wherein said RIZ reduces the growth of said tumor cell.

52. The method of claim 51, wherein said tumor cell is selected from the group consisting of a neuroblastoma tumor cell and a melanoma tumor cell.

53. A method for detecting the presence of a RIZ in a test sample, comprising the steps of:

a. obtaining the test sample;

b. contacting said test sample with a RIZ binding reagent under suitable conditions, which allow specific binding of said reagent to the RIZ; and

c. detecting said specifically bound reagent, which indicates the presence of said RIZ.

54. A method for diagnosing a pathology characterized by an abnormal level of expression of a RIZ in a subject, comprising the method of claim 53, further comprising the steps of:

d. measuring the level of said RIZ; and

e. comparing said level to the level of RIZ in a control normal sample, wherein a significant difference in said levels is diagnostic of a pathology characterized by abnormal level of expression of a RIZ.

55. A method for diagnosing a pathology characterized by expression of a mutant RIZ in a subject, comprising the step of:

a. obtaining the test sample;

b. contacting said test sample with a mutant RIZ binding reagent under suitable conditions, which allow specific binding of said reagent to the mutant RIZ; and

c. detecting said specifically bound reagent, which indicates the presence of said mutant RIZ.

56. A method for detecting the presence of a nucleic acid molecule encoding a RIZ in a test sample, comprising the steps of:

a. obtaining the test sample;

b. contacting said test sample with the nucleotide sequence of claim 9 under suitable conditions, which allow specific binding of said nucleotide sequence to a nucleic acid molecule encoding a RIZ; and

c. detecting said specifically bound nucleotide sequence, which indicates the presence of said nucleic acid molecule encoding said RIZ.

57. A method for diagnosing a pathology characterized by an abnormal level of expression of a nucleic acid molecule encoding a RIZ in a subject, comprising the method of claim 56, further comprising the steps of:

d. measuring the level of said nucleic acid molecule encoding the RIZ; and

e. comparing said level to the level of RIZ in a control normal sample, wherein a significant difference in said levels is diagnostic of a pathology characterized by abnormal level of expression of a RIZ.

58. A method for diagnosing a pathology characterized by a mutant nucleic acid molecule encoding a RIZ in a subject, comprising the steps of:

a. obtaining the test sample;

b. contacting said test sample with the nucleotide sequence of claim 15 under suitable conditions, which allow specific binding of said nucleotide sequence to a mutant nucleic molecule encoding a RIZ; and

c. detecting said specifically bound nucleotide sequence, which indicates the presence of said mutant nucleic acid molecule encoding a RIZ.

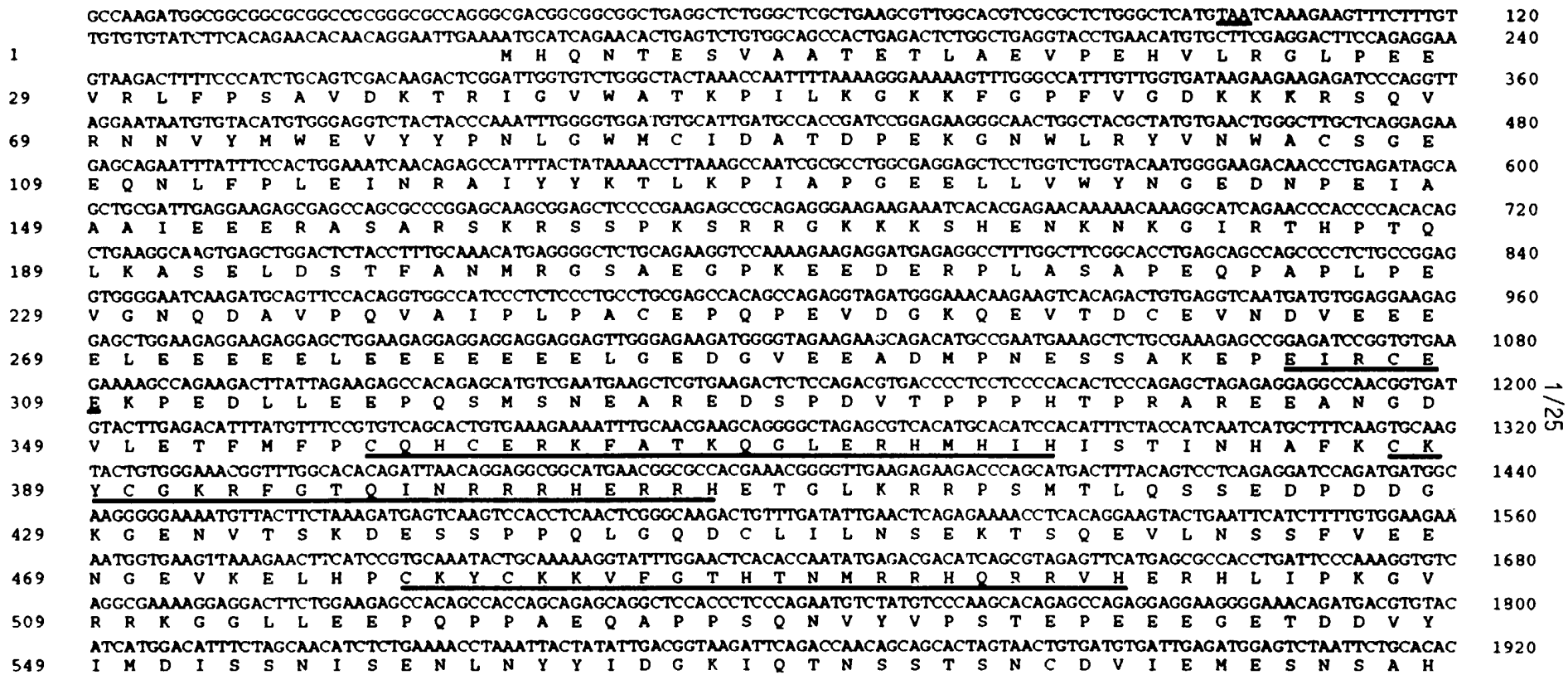

FIG. 1A-1

```
                TTGTATGGCATAGACTGTCTGCTCACTCCAGTGACCGTGGAGATTACTCAGAACATAAAGAGCACTCAGGTCTCTGTGACAGATGATCTTCTCAAAGACTCTCCCAGCAGCACAAATTGT  2040
589             L  Y  G  I  D  C  L  L  T  P  V  T  V  E  I  T  Q  N  I  K  S  T  Q  V  S  V  T  D  D  L  L  K  D  S  P  S  S  T  N  C
                GAGTCTAAGAAACGGAGGACTGCCAGTCCACCTGTGCTCCCCAAAATTAAAACGGAGACGGAGTCTGATTCCACAGCACCCTCGTGTTCCTTAAGTCTGCCCCTGAGCATATCCACAGCC  2160
629             E  S  K  K  R  R  T  A  S  P  P  V  L  P  K  I  K  T  E  T  E  S  D  S  T  A  P  S  C  S  L  S  L  P  L  S  I  S  T  A
                GAGGTGGTGTCCTTCCATAAAGAGAAGGGCGTCTATTTGTCGTCCAAGCTCAAGCAGCTTCTTCAGACCCAGGACAAGCTGACCCTTCCTGCAGGGTTTTCAGCAGCTGAGATTCCTAAG  2280
669             E  V  V  S  F  H  K  E  K  G  V  Y  L  S  S  K  L  K  Q  L  L  Q  T  Q  D  K  L  T  L  P  A  G  F  S  A  A  E  I  P  K
                TTAGGTCCCGTGTGCGCGTCTGCTCCTGCATCCATGTTGCCCGTGACCTCTAGTAGGTTTAAGAGACGCACCAGCTCTCCACCGAGCTCTCCACACGCACAGCCCTGCCCTTCGAGACTTC  2400
709             L  G  P  V  C  A  S  A  P  A  S  M  L  P  V  T  S  S  R  F  K  R  R  T  S  S  P  P  S  S  P  Q  H  S  P  A  L  R  D  F
                GGGAAACCAAATGATGGGAAAGCAGCATGGACAGACACAGTCCTGACTTCCAAGAAACCCAAGTTAGAAAGTCGTAGTGACTCACCAGCATGGAGTTTGTCTGGGAGAGATGAAAGAGAA  2520
749             G  K  P  N  D  G  K  A  A  W  T  D  T  V  L  T  S  K  K  P  K  L  E  S  R  S  D  S  P  A  W  S  L  S  G  R  D  E  R  E
                ACCGGAAGCCCTCCTTGCTTTGATGAATACAAAATATCAAAGGAATGGGCAGCCAGTTCTACTTTCAGCAGTGTGTGCAACCAACAGCCATTGGATTTATCCAGCGGGGTCAAACAGAAG  2640
789             T  G  S  P  P  C  F  D  E  Y  K  I  S  K  E  W  A  A  S  S  T  F  S  S  V  C  N  Q  Q  P  L  D  L  S  S  G  V  K  Q  K
                TCAGAGGGCACAGGCAAGACTCCAGTCCCATGGGAATCTGTATTGGATCTCAGTGTGCATAAAAAGCCTTGCGATTCTGAAGGCAAGGAATTCAAAGAGAACCATTTGGCACAGCCAGCT  2760
829             S  E  G  T  G  K  T  P  V  P  W  E  S  V  L  D  L  S  V  H  K  K  P  C  D  S  E  G  K  E  F  K  E  N  H  L  A  Q  P  A
                GCAAAGAAGAAAAAACCAACCACCTGTATGCTTCAAAAGGTTCTTCTCAATGAGTATAATGGTGTTAGCTTACCTACAGAAACCACACCAGAGGTGACCAGGAGCCCAAGTCCTTGTAAA  2880
869             A  K  K  K  K  P  T  T  C  M  L  Q  K  V  L  L  N  E  Y  N  G  V  S  L  P  T  E  T  T  P  E  V  T  R  S  P  S  P  C  K
                TCCCCAGATACACAGCCAGATCCTGAACTTGGTCCTGACTCAAGTTGCTCAGTCCCCACTGCTGAGTCTCCACCTGAAGTTGTTGGCCCTTCCTCACCCCCTCTCCAGACAGCCTCCTTA  3000
909             S  P  D  T  Q  P  D  P  E  L  G  P  D  S  S  C  S  V  P  T  A  E  S  P  P  E  V  V  G  P  S  S  P  P  L  Q  T  A  S  L
                TCCTCCGGTCAGCTGCCTCCTCTCTTAACCCCCACAGAGCCTTCTTCCCCTCCCCCCTGCCCTCCTGTGTTAACTGTTGCCACTCCACCACCTCCCCTCCTTCCAACCGTCCCTCTCTCC  3120
949             S  S  G  Q  L  P  P  L  L  T  P  T  E  P  S  S  P  P  P  C  P  P  V  L  T  V  A  T  P  P  P  P  L  L  P  T  V  P  L  S
                CACCCCTCTTCTGATGCCTCCCCTCAGCAGTGTCCCTCTCCGTTCTCAAACACCACTGCTCAGTCTCCTCTTCCCATTCTCTCCCCCAACAGTGTCTCCCTCTCCCTCTCCCATTCCTCCT  3240
989             H  P  S  S  D  A  S  P  Q  Q  C  P  S  P  F  S  N  T  T  A  Q  S  P  L  P  I  L  S  P  T  V  S  P  S  P  S  P  I  P  P
                GTAGAGCCACTTATGTCTGCTGCTTCCCCTGGTCCCCCAACACTTTCTTCCTCCTCCTCTTCTTCCTCTTCCTTCCCTTCCTCTTCCTGCTCCTCCACCTCCCCCTCCCCACCCCCTCTT  3360
1029            V  E  P  L  M  S  A  A  S  P  G  P  P  T  L  S  S  S  S  S  S  S  S  S  F  P  S  S  S  C  S  S  T  S  P  S  P  P  P  L
                TCAGCAGTGTCATCTGTGGTTTCCTCTGGGGACAACCTGGAGGCATCTCTGCCTGCAGTAACTTTCAAACAGGAGGAGTCAGAGAGTGAAGGTCTGAAACCCAAGGAAGAGGCCCCACCT  3480
1069            S  A  V  S  S  V  V  S  S  G  D  N  L  E  A  S  L  P  A  V  T  F  K  Q  E  E  S  E  S  E  G  L  K  P  K  E  E  A  P  P
                GCAGGGGGACAGAGTGTGGTCCAAGAAACATTCAGCAAAAACTTCATTTGCAATGTCTGTGAATCGCCTTTTCTTTCCATTAAAGACCTAACCAAACATTTATCCGTCCATGCTGAAGAG  3600
1109            A  G  G  Q  S  V  V  Q  E  T  F  S  K  N  F  I  C  N  V  C  E  S  P  F  L  S  I  K  D  L  T  K  H  L  S  V  H  A  E  E
                TGGCCCTTCAAATGTGAGTTTTGTGTGCAGCTGTTTAAGGTTAAGACTGATCTATCAGAGCATCGATTTCTGCTTCATGGGGTTGGAAATATCTTTGTGTGTTCTGTATGTAAGAAAGAA  3720
1149            W  P  F  K  C  E  F  C  V  Q  L  F  K  V  K  T  D  L  S  E  H  R  F  L  L  H  G  V  G  N  I  F  V  C  S  V  C  K  K  E
                TTTGCCTTCTTATGCAATCTGCAGCAGCACCAGCGTGATCTCCACCCAGATGAGGTATGCACACACCACGAGTTTGAAAGTGGGACCCTGAGGCCCCAGAACTTCACAGACCCCAGCAAG  3840
1189            F  A  F  L  C  N  L  Q  Q  H  Q  R  D  L  H  P  D  E  V  C  T  H  H  E  F  E  S  G  T  L  R  P  Q  N  F  T  D  P  S  K
```

FIG. 1A-2

```
     GCCAATGTTGAGCATATGCCAAGTTTGCCAGAAGAGCCTTTAGAAACTTCTAGAGAGGAGGAGTTAAATGATTCCTCTGAAGAGCTTTACACGACCATCAAAATAATGGCTTCTGGAATA  3960
1229 A  N  V  E  H  M  P  S  L  P  E  E  P  L  E  T  S  R  E  E  E  L  N  D  S  S  E  E  L  Y  T  T  I  K  I  M  A  S  G  I
     AAGACGAAGGATCCAGATGTTCGACTTGGTCTCAACCAGCACTACCCGAGCTTTAAACCTCCTCCATTTCAGTACCACCATCGAAACCCTATGGGGATAGGGGTGACAGCCACCAACTTC  4080
1269 K  T  K  D  P  D  V  R  L  G  L  N  Q  H  Y  P  S  F  K  P  P  P  F  Q  Y  H  H  R  N  P  M  G  I  G  V  T  A  T  N  F
     ACCACCCACAATATTCCACAGACTTTCACTACTGCCATCCGCTGCACAAAGTGTGGGAAGGGCGTCGACAATATGCCTGAGCTGCATAAGCATATCTTGGCGTGTGCGTCTGCAAGTGAC  4200
1309 T  T  H  N  I  P  Q  T  F  T  T  A  I  R  C  T  K  C  G  K  G  V  D  N  M  P  E  L  H  K  H  I  L  A  C  A  S  A  S  D
     AAGAAGAGGTACACCCCTAAGAAAAACCCAGTGCCCCTGAAACAAACTGTGCAGCCCAAAAATGGAGTGGTGGTTCTAGACAACTCTGGGAAAAATGCCTTCAGACGGATGGGGCAGCCC  4320
1349 K  K  R  Y  T  P  K  K  N  P  V  P  L  K  Q  T  V  Q  P  K  N  G  V  V  V  L  D  N  S  G  K  N  A  F  R  R  M  G  Q  P
     AAGAGACTGAGCTTCAATGTTGAACTGGGTAAAATGTCTCCAAACAAGCTCAAGCTGAGTGCGCTGAAGAAGAAAAACCAGCTGGTGCAGAAGGCCATCCTTCAGAAGAACAGAGCCGCG  4440
1389 K  R  L  S  F  N  V  E  L  G  K  M  S  P  N  K  L  K  L  S  A  L  K  K  K  N  Q  L  V  Q  K  A  I  L  Q  K  N  R  A  A
     AAGCAGAAGGCGGACCTGAGGGATACTTCCGAGGCGTCCTCACACATCTGCCCGTACTGTGACAGGGAGTTCACATACATTGGCAGCCTGAATAAGCATGCCGCCTTCAGCTGTCCTAAA  4560
1429 K  Q  K  A  D  L  R  D  T  S  E  A  S  S  H  I  C  P  Y  C  D  R  E  F  T  Y  I  G  S  L  N  K  H  A  A  F  S  C  P  K
     AAACCTCTTTCTCCTTCCAAAAGAAAAGTTTCCCATTCGTCTAAGAAAGGTGGCCATGCATCATCTTCTAGCAGTGACAGAAACAGCAGCTGCCACCCCCGGAGGCGGACTGCAGATACC  4680
1469 K  P  L  S  P  S  K  R  K  V  S  H  S  S  K  K  G  G  H  A  S  S  S  S  S  D  R  N  S  S  C  H  P  R  R  R  T  A  D  T
     GAGATCAAGATGCAGAGCACGCAGGCACCCTTGGGCAAGACCAGAGCTCGGAGTACAGGCCCCGCCCAGGCCTCACTGCCCTCCTCGTCCTTCAGATCCAGACAGAATGTCAAATTTGCA  4800
1509 E  I  K  M  Q  S  T  Q  A  P  L  G  K  T  R  A  R  S  T  G  P  A  Q  A  S  L  P  S  S  S  F  R  S  R  Q  N  V  K  F  A
     GCTTCAGTGAAATCCAAAAAAGCAAGCTCTTCATCCTTGAGGAATTCCAGTCCCATAAGAATGGCCAAAATTACTCACGTCGAGGGCAAAAAACCCAAAGCTGTTGCCAAGAGTCATTCT  4920
1549 A  S  V  K  S  K  K  A  S  S  S  S  L  R  N  S  S  P  I  R  M  A  K  I  T  H  V  E  G  K  K  P  K  A  V  A  K  S  H  S
     GCTCAGCTCTCAAGCAAATCCTCCCGAGGCCTGCATGTCAGAGTGCAGAAGAGCAAAGCTGTCATACAGAGCAAGACTGCCCTGGCCAGTAAGAGGAGAACAGACCGGTTCATAGTGAAA  5040
1589 A  Q  L  S  S  K  S  S  R  G  L  H  V  R  V  Q  K  S  K  A  V  I  Q  S  K  T  A  L  A  S  K  R  R  T  D  R  F  I  V  K
     TCTAGAGAGCGCAGCGGGGGCCCAATCACCCGAAGCCTTCAGCTGGCAGCTGCTGCGGACCTGAGTGAAAGCAGGAGAGAGGACAGCAGTGCCAGGCATGAGCTGAAGGACTTCAGCTAC  5160
1629 S  R  E  R  S  G  G  P  I  T  R  S  L  Q  L  A  A  A  A  D  L  S  E  S  R  R  E  D  S  S  A  R  H  E  L  K  D  F  S  Y
     AGTCTCCGCCTGGCATCTCGATGCGGCTCATCAACAGCCTCTTACATCACCAGACAATGCAGAAAGGTCAAGGCCGCCGCAGCAACTCCGTTCCAGGGACCCTTCCTCAAAGAGTAGGCA  5280
1669 S  L  R  L  A  S  R  C  G  S  S  T  A  S  Y  I  T  R  Q  C  R  K  V  K  A  A  A  A  T  P  F  Q  G  P  F  L  K  E
     CTCTGTCTGCTCCTTAACAGCACCTGAAGTGACCTGGAATCAGTGAAGCCAAAGGGACCAGCAGTCTGCCCTGCAGAGAGCACTGACCTCTCCCAGTTGTGAGAGTGAGAGAACGAGAGA  5400
     GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAATGAGAATGTGTGTGTGTGTGTGCTGGTGCATGTGTGTGGTCTTCAAGCCAAGGTCCCAGCCTCAGGAGCAGGACCTTCCCA  5520
     TTTCCCGTCATCCTCTGGATGATCCTTGGACGTGGCCCAGAACCGTGCTCTGTGGTGCAGCCATCCTGCCCGGGAGGGGCATCTCCTTCTATGCAATTTTTTTAAAGAGTTCCTTGGCCC  5640
     TGCTTTGTGCTTCTTGAGCTGTCCGTTTGCCACCACTGGGACTTGGATCTGGCCCTGAGGGGTGGGGAAGAGGGCCTATCTAAGGATAACCTTTCAGAGGTCAAGCTCCCCTTCATGCCA  5760
     CCCCTCCCCCCTGCCCTCACCGACCTTTTCCCCACACTGTCTCTGGGAATCAATAGCAGATAGCATATAGATCCATCAGGGTTGAGCCTGAACCTCGGCCCTAGCACTAGGAAATCCCCC  5880
     TTTTCTCCCTAAGCAACTGGAGCCGCCAGCTTTCAAGTCATTTCCTCCTTTGAGGTTCTAGAGTCCGAGAGTCTGCTCCGAAGTCTCTCCTGGGAACCCGGGAGCCCTCGCACCCAGGAC  6000
     GCAGACTCTGTGCCCATTCTTAGACCTGAGGTAGAAGAAGCAGTGTTTTGGGACGATAGGGTGGAGGCGTGCCTACTTTGTCTCCTCTGGTGGGACCTCCTACATCATTGGCATCTGAAC  6120
     CTTGCAAGTTCGCTGCAAAGAGAAGCAAAGGAAAAAAAAAAAAAAAAAAAA                                                                        6171
```

FIG. 1A–3

GAATTCCCGG CTCACTGAAG CTTGGCACGT GCGCTCTGGA

ATATCTGAAT GATCTCAGTA CAATGAAGGA GTGCCTTTTC

CCTTTCTACC CTGCCTCCTT GAAGCATGCA TTAGAGTCGT

T (SEQ ID NO: 94)

FIG. 1B

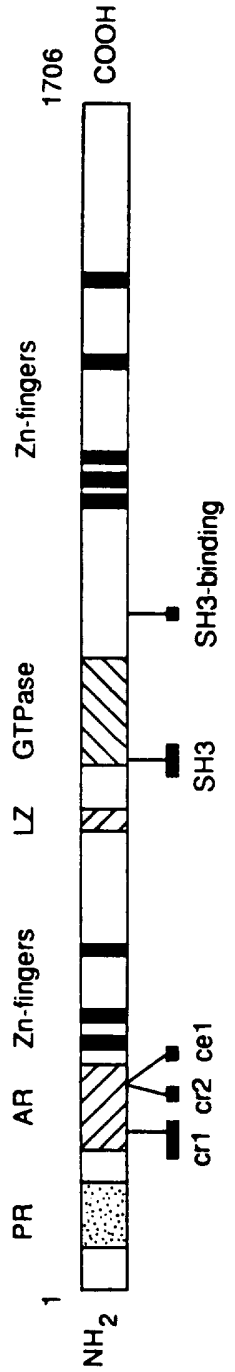

FIG. 3

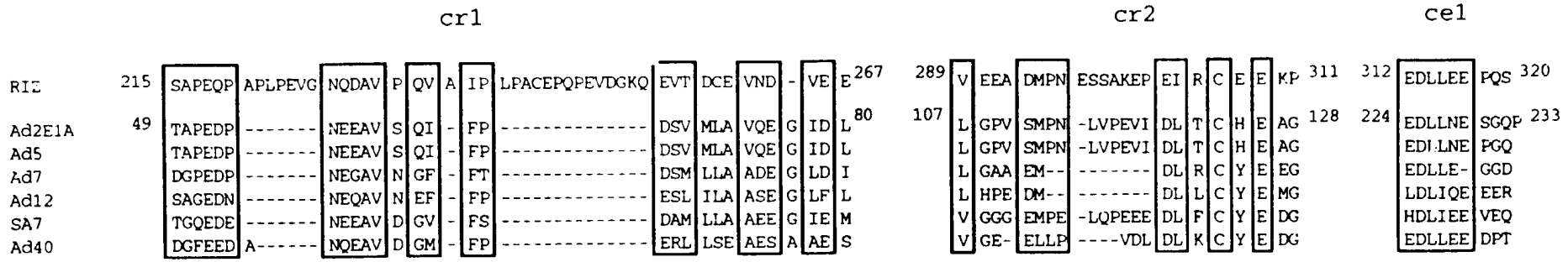

FIG. 2A

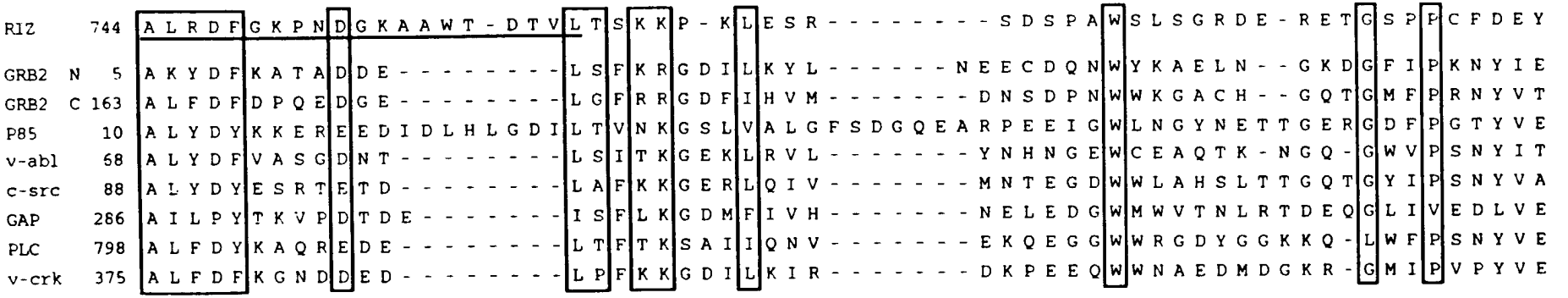

FIG. 2B-1

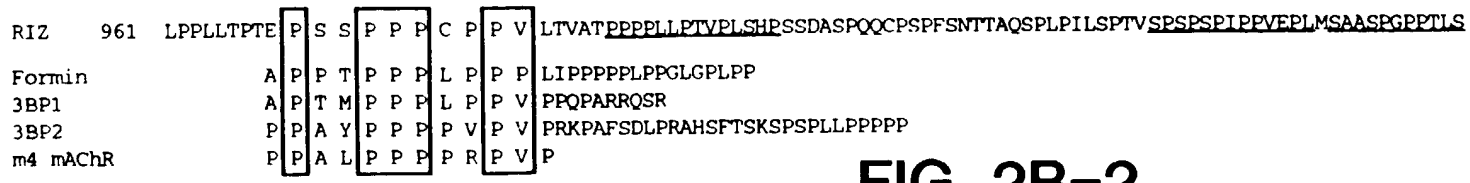

FIG. 2B-2

```
RIZ          1125  CNVCESPFLSIKDLTKHLSVHAEEWPFKCEFCVQLFKVKTDLSEHRFLLHGVGNIFVCSVCKKEFAFLCNLQQHQRDLHPDEVCTH  1210
PRDI-BF1      543  CNVCAKTFGQLSNLKVHLRVHSGERPFKCQTCNKGFTQLAHLQKH-YLVHTGEKPHECQVCHKRFSSTSNLKTHLR-LHSGEKPYQ  626
CONSENSUS          CNVC...F.....L..HL.VH..E.PFKC..C...F.....L..H..L.H.......C.VC.K.F....NL..H.R.LH..E....
DNA-CONTACTS             *  *  *                    *  *  *                    *  *  *
```

FIG. 2C-1

```
RIZ            39  TRIGVWATKPILKGKKFGPFVGDKKKRSQVRNNV---YMWEVYYPNLGWMCIDATDPEKGNWLRYVNWACSGEEQNLFPL  115
PRDI-BF1       60  EVIGVMSKEYIPKGTRFGPLIGEIYTNDTVPKNANRKYFWRIYSRGELHHFIDGFNEEKSNWMRYVNPAHSPREQNLAAC  139
CONSENSUS          ..IGV.....I.KG..FGP..G.......V..N....Y.W..Y........ID....EK.NW.RYVN.A.S..EQNL...

RIZ            116  EINRAIYYKTLKPIAPGEELLVWYNGEDNP  145
PRDI-BF1       140  QNGMNIYFYTIKPIPANQELLVWYCRDFAE  169
CONSENSUS          .....IY..T.KPI....ELLVWY......
```

FIG. 2C-2

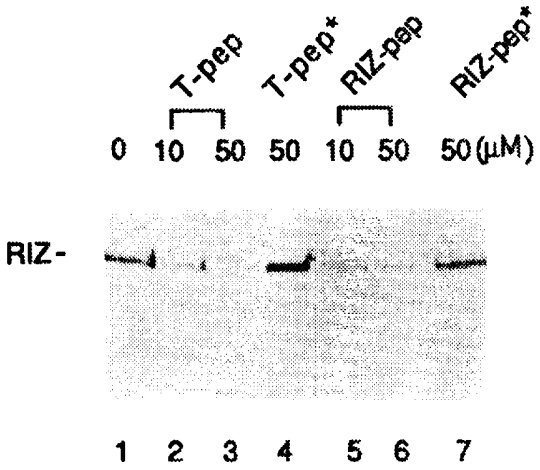

FIG. 4

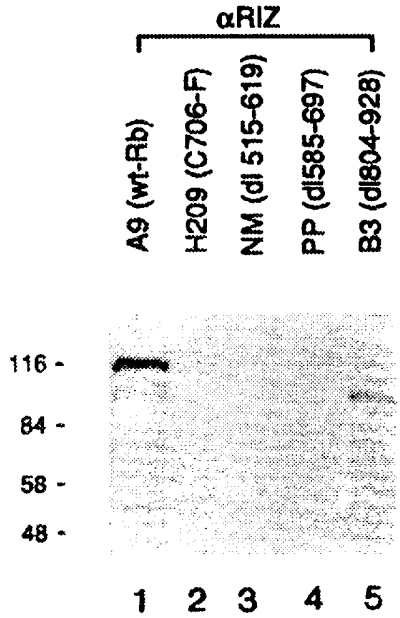

FIG. 5A

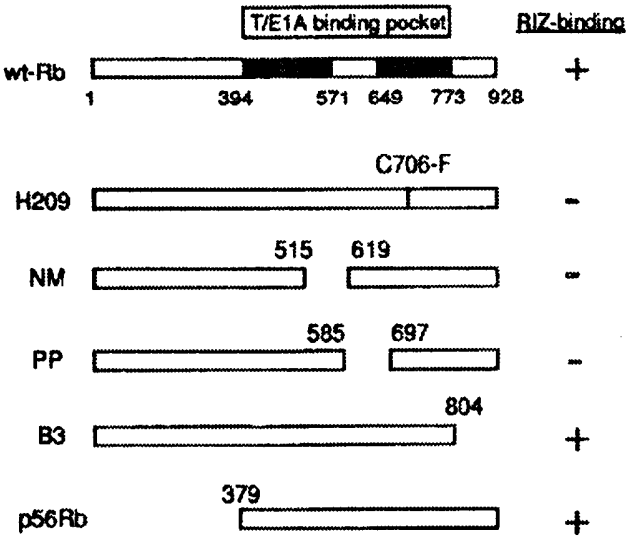

FIG. 5B

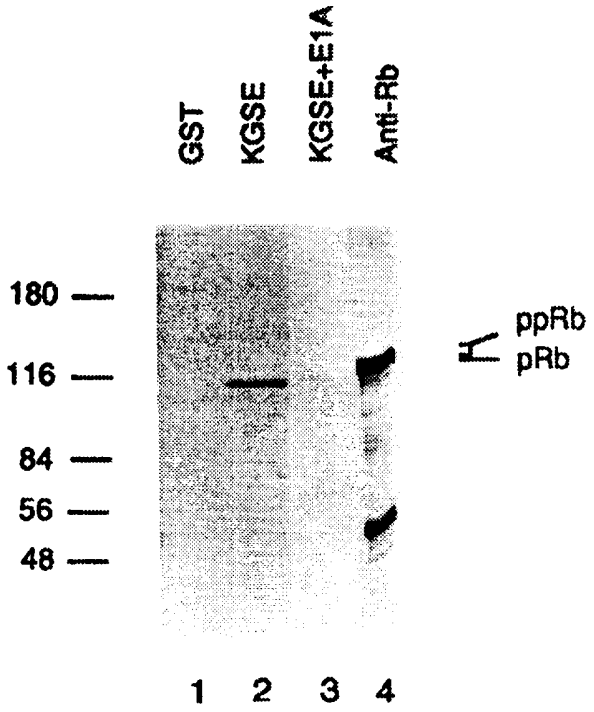

FIG. 6

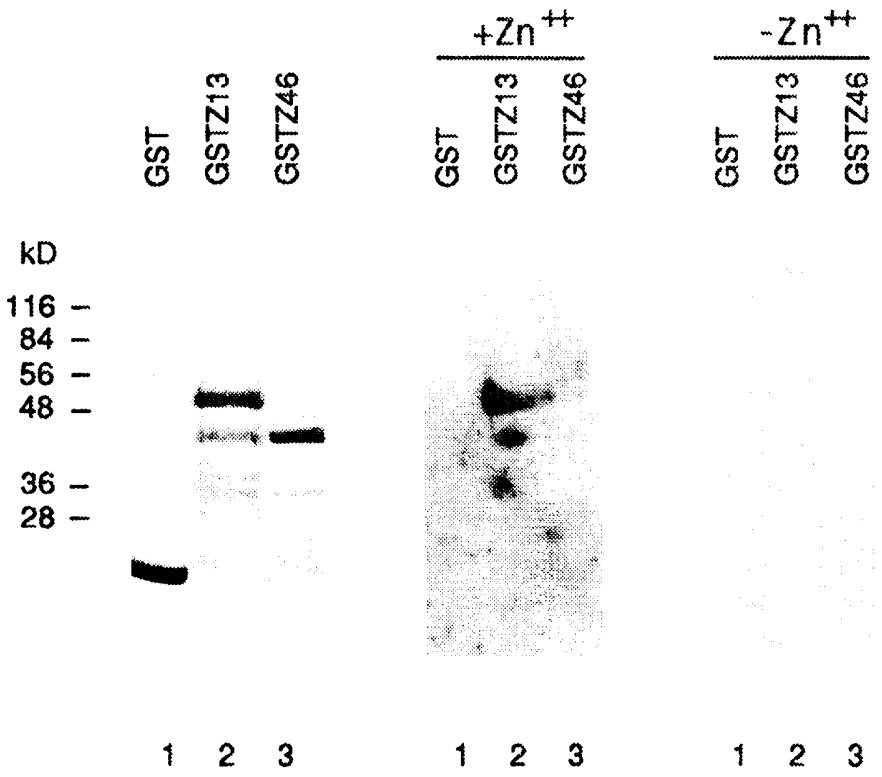

FIG. 7A FIG. 7B FIG. 7C

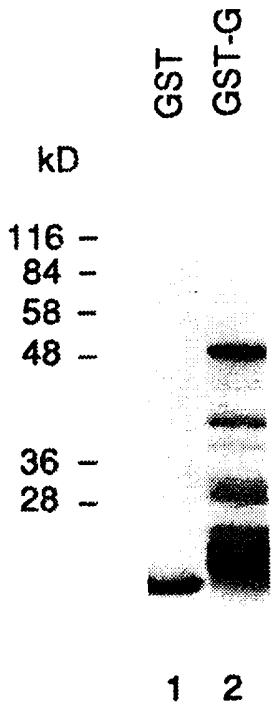

FIG. 8A

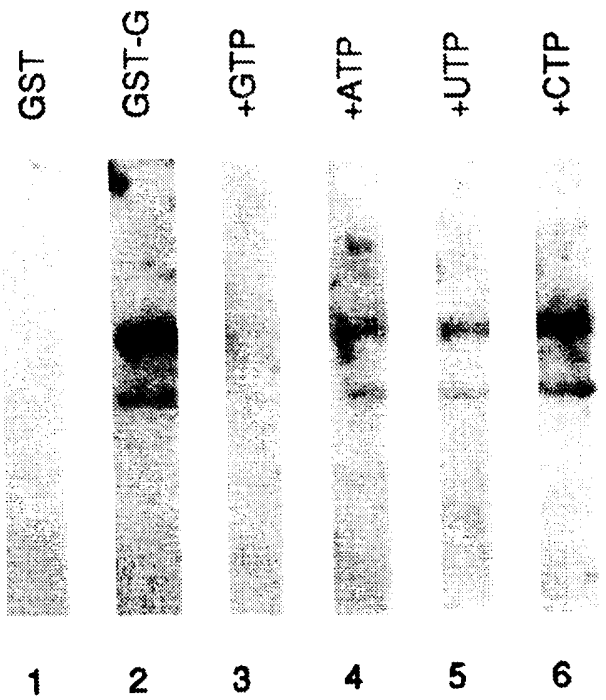

FIG. 8B

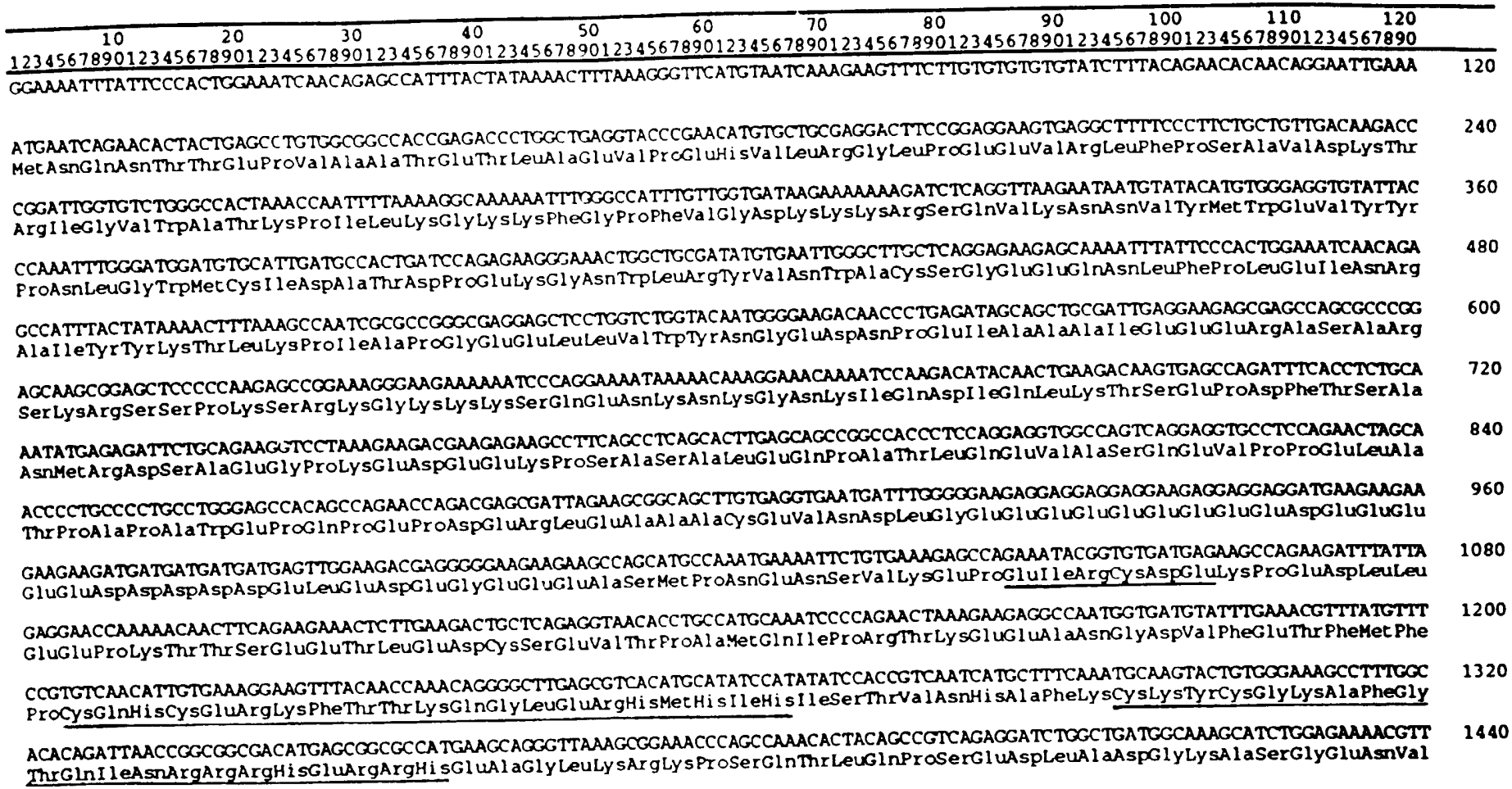

FIG. 9A-1

```
GCTTCAAAAGATGATTCGAGTCCTCCCAGTCTTGGGCCAGACTGTCTGATCATGAATTCAGAGAAGGCTTCCCAAGACACAATAAATTCTTCTGTCGTAGAAGAGAATGGGGAAGTTAAA  1560
AlaSerLysAspAspSerSerProProSerLeuGlyProAspCysLeuIleMetAsnSerGluLysAlaSerGlnAspThrIleAsnSerSerValValGluGluAsnGlyGluValLys

GAACTTCATCCGTGCAAATATTGTAAAAAGGTTTTTGGAACTCATACTAATATGAGACGGCATCAGCGTAGAGTTCACGAACGTCATCTGATTCCCAAAGGTGTACGGCGAAAAGGAGGC  1680
GluLeuHisProCysLysTyrCysLysLysValPheGlyThrHisThrAsnMetArgArgHisGlnArgArgValHisGluArgHisLeuIleProLysGlyValArgArgLysGlyGly

CTTGAAGAGCCCCAGCCTCCAGCAGAACAGGCCCAGGCCACCCAGAACGTGTATGTACCAAGCACAGAGCCGGAGGAGGAAGGGGAAGCAGATGATGTGTACATCATGGACATTTCTAGC  1800
LeuGluGluProGlnProProAlaGluGlnAlaGlnAlaThrGlnAsnValTyrValProSerThrGluProGluGluGluGlyGluAlaAspAspValTyrIleMetAspIleSerSer

AATATCTCTGAAAACTTAAATTACTATATTGATGGTAAAATTCAAACTAATAACAACACTAGTAACTGTGATGTGATTGAGATGGAGTCTGCTTCGGCAGATTTGTATGGTATAAATTGT  1920
AsnIleSerGluAsnLeuAsnTyrTyrIleAspGlyLysIleGlnThrAsnAsnAsnThrSerAsnCysAspValIleGluMetGluSerAlaSerAlaAspLeuTyrGlyIleAsnCys

CTGCTCACTCCAGTTACAGTGGAAATTACTCAAAATATAAAGACCACACAGGTCCCTGTAACAGAAGATCTTCCTAAAGAGCCTTTGGGCAGCACAAATAGTGAGGCCAAGAAGCGGAGA  2040
LeuLeuThrProValThrValGluIleThrGlnAsnIleLysThrThrGlnValProValThrGluAspLeuProLysGluProLeuGlySerThrAsnSerGluAlaLysLysArgArg

ACTGCGAGCCCACCTGCACTGCCCAAAATTAAGGCCGAAACAGACTCTGACCCCATGGTCCCCTCTTGCTCTTTAAGTCTTCCTCTTAGCATATCAACAACAGAGGCAGTGTCTTTCCAC  2160
ThrAlaSerProProAlaLeuProLysIleLysAlaGluThrAspSerAspProMetValProSerCysSerLeuSerLeuProLeuSerIleSerThrThrGluAlaValSerPheHis

AAAGAGAAAAGTGTTTATTTGTCATCAAAGCTCAAACAACTTCTTCAAACCCAAGATAAACTAACTCCTCCTGCAGGGATTTCAGCAACTGAAATAGCTAAATTAGGTCCTGTTTGTGTG  2280
LysGluLysSerValTyrLeuSerSerLysLeuLysGlnLeuLeuGlnThrGlnAspLysLeuThrProProAlaGlyIleSerAlaThrGluIleAlaLysLeuGlyProValCysVal

TCTGCTCCTGCATCAATGTTGCCTGTGACCTCAAGTAGGTTTAAGAGGCGGACCAGCTCTCCTCCCAGTTCTCCACAGCACAGTCCTGCCCTTCGAGACTTTGGAAAGCCAAGTGATGGG  2400
SerAlaProAlaSerMetLeuProValThrSerSerArgPheLysArgArgThrSerSerProProSerSerProGlnHisSerProAlaLeuArgAspPheGlyLysProSerAspGly

AAAGCAGCATGGACCGATGCCGGGCTGACTTCCAAAAAATCCAAATTAGAAAGTCACAGCGACTCACCAGCATGGAGTTTGTCTGGGAGAGATGAGAGAGAAACTGTGAGCCCTCCATGC  2520
LysAlaAlaTrpThrAspAlaGlyLeuThrSerLysLysSerLysLeuGluSerHisSerAspSerProAlaTrpSerLeuSerGlyArgAspGluArgGluThrValSerProProCys

TTTGATGAATATAAAAATGTCTAAAGAGTGGACAGCTAGTTCTGCTTTTAGCAGTGTGTGCAACCAGCAGCCACTGGATTTATCCAGCGGTGTCAAACAGAAGGCTGAGGGTACAGGCAAG  2640
PheAspGluTyrLysMetSerLysGluTrpThrAlaSerSerAlaPheSerSerValCysAsnGlnGlnProLeuAspLeuSerSerGlyValLysGlnLysAlaGluGlyThrGlyLys

ACTCCAGTCCAGTGGGAATCTGTCTTAGATCTCAGTGTGCATAAAAAGCATTGTAGTGACTCTGAAGGCAAGGAATTCAAAGAAAGTCATTCAGTGCAGCCTACGTGTAGTGCTGTAAAG  2760
ThrProValGlnTrpGluSerValLeuAspLeuSerValHisLysLysHisCysSerAspSerGluGlyLysGluPheLysGluSerHisSerValGlnProThrCysSerAlaValLys

AAAAGGAAACCAACCACCTGCATGCTGCAGAAGGTTCTTCTCAATGAATATAATGGCATCGATTTACCTGTAGAAAACCCTGCAGATGGGACCAGGAGCCCAAGTCCTTGTAAATCCCTA  2880
LysArgLysProThrThrCysMetLeuGlnLysValLeuLeuAsnGluTyrAsnGlyIleAspLeuProValGluAsnProAlaAspGlyThrArgSerProSerProCysLysSerLeu
```

FIG. 9A-2

```
GAAGCTCAGCCAGATCCTGACCTCGGTCCGGGCTCTGGTTTCCCTGCCCCTACTGTTGAGTCCACACCTGATGTTTGTCCTTCATCACCTGCCCTGCAGACACCCTCCCTTTCATCCGGT  3000
GluAlaGlnProAspProAspLeuGlyProGlySerGlyPheProAlaProThrValGluSerThrProAspValCysProSerSerProAlaLeuGlnThrProSerLeuSerSerGly

CAGCTGCCTCCTCTCTTGATCCCCACAGATCCCTCTTCCCCTCCACCCTGTCCCCCGGTATTAACTGTTGCCACTCCGCCCCCTCCCCTCCTTCCTACCGTACCTCTTCCAGCCCCCTCT  3120
GlnLeuProProLeuLeuIleProThrAspProSerSerProProProCysProProValLeuThrValAlaThrProProProProLeuLeuProThrValProLeuProAlaProSer

TCCAGTGCATCTCCACACCCATGCCCCTCTCCACTCTCAAATGCCACCGCACAGTCCCCACTTCCAATTCTGTCCCCAACAGTGTCCCCCCTCTCCCTCTCCCATTCCTCCCCGTGGAGCCC  3240
SerSerAlaSerProHisProCysProSerProLeuSerAsnAlaThrAlaGlnSerProLeuProIleLeuSerProThrValSerProSerProSerProIleProProValGluPro

CTGATGTCTGCCGCCTCACCCGGGCCTCCAACACTTTCTTCTTCCTCCTCTTCATCTTCCTCCTCCTCTTCGTTTTCTTCTTCATCTTCCTCCTCTTCTCCTTCTCCACCTCCTCTCTCC  3360
LeuMetSerAlaAlaSerProGlyProProThrLeuSerSerSerSerSerSerSerSerSerSerSerSerPheSerSerSerSerSerSerSerSerProSerProProProLeuSer

GCAATATCATCTGTTGTTTCCTCTGGTGATAATCTGGAGGCTTCTCTCCCCATGATATCTTTCAAACAGGAGGAATTAGAGAATGAAGGTCTGAAACCCAGGGAAGAGCCCCAGTCTGCT  3480
AlaIleSerSerValValSerSerGlyAspAsnLeuGluAlaSerLeuProMetIleSerPheLysGlnGluGluLeuGluAsnGluGlyLeuLysProArgGluGluProGlnSerAla

GCTGAACAGGATGTTGTTGTTCAGGAAACATTCAACAAAAACTTTGTTTGCAACGTCTGTGAATCACCTTTTCTTTCCATTAAAGATCTAACCAAACATTTATCTATTCATGCTGAAGAA  3600
AlaGluGlnAspValValValGlnGluThrPheAsnLysAsnPheValCysAsnValCysGluSerProPheLeuSerIleLysAspLeuThrLysHisLeuSerIleHisAlaGluGlu

        10        20        30        40        50        60        70        80        90       100       110       120
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789

TGGCCCTTCAAATGTGAATTTTGTGTGCAGCTTTTTAAGGATAAAACGGACTTGTCAGAACATCGCTTTTTGCTTCATGGAGTTGGGAATATCTTTGTGTGTTCTGTTTGTAAAAAAGAA  3720
TrpProPheLysCysGluPheCysValGlnLeuPheLysAspLysThrAspLeuSerGluHisArgPheLeuLeuHisGlyValGlyAsnIlePheValCysSerValCysLysLysGlu

TTTGCTTTTTTGTGCAATTTGCAGCAGCACCAGCGAGATCTCCACCCAGATAAGGTGTGCACACATCACGAGTTTGAAAGCGGGACTCTGAGGCCCCAGAACTTTACAGATCCCAGCAAG  3840
PheAlaPheLeuCysAsnLeuGlnGlnHisGlnArgAspLeuHisProAspLysValCysThrHisHisGluPheGluSerGlyThrLeuArgProGlnAsnPheThrAspProSerLys

GCCCATGTAGAGCATATGCAGAGCTTGCCAGAAGATCCTTTAGAAACTTCTAAAGAAGAAGAGGAGTTAAATGATTCCTCTGAAGAGCTTTACACGACTATAAAAATAATGGCTTCTGGA  3960
AlaHisValGluHisMetGlnSerLeuProGluAspProLeuGluThrSerLysGluGluGluGluLeuAsnAspSerSerGluGluLeuTyrThrThrIleLysIleMetAlaSerGly

ATAAAGACAAAAGATCCAGATGTTCGATTGGGCCTCAATCAGCATTACCCAAGCTTTAAACCACCTCCATTTCAGTACCATCACCGTAACCCCATGGGGATTGGTGTGACAGCCACAAAT  4080
IleLysThrLysAspProAspValArgLeuGlyLeuAsnGlnHisTyrProSerPheLysProProProPheGlnTyrHisHisArgAsnProMetGlyIleGlyValThrAlaThrAsn

TTCACTACACACAATATTCCACAGACTTTCACTACCGCCATTCGCTGCACAAAGTGTGGAAAAGGTGTCGACAATATGCCGGAGTTGCACAAACATATCCTGGCTTGTGCTTCTGCAAGT  4200
PheThrThrHisAsnIleProGlnThrPheThrThrAlaIleArgCysThrLysCysGlyLysGlyValAspAsnMetProGluLeuHisLysHisIleLeuAlaCysAlaSerAlaSer

GACAAGAAGAGGTACACGCCTAAGAAAAACCCAGTACCATTAAAACAAACTGTGCAACCCAAAAATGGCGTGGTGGTTTTAGATAACTCTGGGAAAAATGCCTTCCGACGAATGGGACAG  4320
AspLysLysArgTyrThrProLysLysAsnProValProLeuLysGlnThrValGlnProLysAsnGlyValValValLeuAspAsnSerGlyLysAsnAlaPheArgArgMetGlyGln
```

FIG. 9A-3

CCCAAAAGGCTTAACTTTAGTGTTGAGCTCAGCAAAATGTCGTCGAATAAGCTCAAATTAAATGCATTGAAGAAAAAAAAATCAGCTAGTACAGAAAGCAATTCTTCAGAAAAACAAATCT 4440
ProLysArgLeuAsnPheSerValGluLeuSerLysMetSerSerAsnLysLeuLysLeuAsnAlaLeuLysLysLysAsnGlnLeuValGlnLysAlaIleLeuGlnLysAsnLysSer

GCAAAGCAGAAGGCCGACTTGAAAAATGCTTGTGAGTCATCCTCTCACATCTGCCCTTACTGTAATCGAGAGTTCACTTACATTGGAAGCCTGAATAAACACGCCGCCTTCAGCTGTCCC 4560
AlaLysGlnLysAlaAspLeuLysAsnAlaCysGluSerSerSerHisIleCysProTyrCysAsnArgGluPheThrTyrIleGlySerLeuAsnLysHisAlaAlaPheSerCysPro

AAAAAACCCCTTTCTCCTCCCAAAAAAAAAGTTTCTCATTCATCTAAGAAAGGTGGACACTCATCACCTGCAAGTAGTGACAAAAACAGTAACAGCAACCACCGCAGACGGACAGCGGAT 4680
LysLysProLeuSerProProLysLysLysValSerHisSerSerLysLysGlyGlyHisSerSerProAlaSerSerAspLysAsnSerAsnSerAsnHisArgArgArgThrAlaAsp

GCGGAGATTAAAATGCAAAGCATGCAGACTCCGTTGGGCAAGACCAGAGCCCGCAGCTCAGGCCCCACCCAAGTCCCACTTCCCTCCTCATCCTTCAGGTCCAAGCAGAACGTCAAGTTT 4800
AlaGluIleLysMetGlnSerMetGlnThrProLeuGlyLysThrArgAlaArgSerSerGlyProThrGlnValProLeuProSerSerSerPheArgSerLysGlnAsnValLysPhe

GCAGCTTCGGTGAAATCCAAAAAACCAAGCTCCTCCTCTTTAAGGAACTCCAGCCCGATAAGAATGGCCAAAATAACTCATGTTGAGGGGAAAAAACCTAAAGCTGTGGCCAAGAATCAT 4920
AlaAlaSerValLysSerLysLysProSerSerSerSerLeuArgAsnSerSerProIleArgMetAlaLysIleThrHisValGluGlyLysLysProLysAlaValAlaLysAsnHis

TCTGCTCAGCTTTCCAGCAAAACATCGCGGAGCCTGCACGTGAGGGTACAGAAAAGCAAAGCTGTTTTACAAAGCAAATCCACCTTGGCGAGTAAGAAAAGAACAGACCGGTTCAATATA 5040
SerAlaGlnLeuSerSerLysThrSerArgSerLeuHisValArgValGlnLysSerLysAlaValLeuGlnSerLysSerThrLeuAlaSerLysLysArgThrAspArgPheAsnIle

AAATCTAGAGAGCGGAGTGGGGGGCCAGTCACCCGGAGCCTTCAGCTGGCAGCTGCTGCTGACTTGAGTGAGAACAAGAGAGAGGACGGCAGCGCCAAGCAGGAGCTGAAGGACTTCAGC 5160
LysSerArgGluArgSerGlyGlyProValThrArgSerLeuGlnLeuAlaAlaAlaAlaAspLeuSerGluAsnLysArgGluAspGlySerAlaLysGlnGluLeuLysAspPheSer

TACAGCCTCCGCTTGGCGTCCCGATGCTCTCCACCAGCGGCCCCGTACATCACCAGGCAGTATAGGAAGGTCAAAGCTCCGGCTGCAGCCCAGTTCCAGGGACCATTCTTCAAAGAGTAG 5280
TyrSerLeuArgLeuAlaSerArgCysSerProProAlaAlaProTyrIleThrArgGlnTyrArgLysValLysAlaProAlaAlaAlaGlnPheGlnGlyProPhePheLysGlu

ACACTCTGGCTGCTCCCTGACAGCACCTGAAGTGACCTGGAATCAGTGAAGCCAAAGGGACTGGCAGTCTGCCTGCAGGGAGTACCGACCTATCCCAGTTGTGTGAGGCTGCGAGAGAAA 5400

GGGAGTGCATGTGCGCGCGTGCATGTGTGCGTGCGTGTGTGTTCACGTGTTCTCGTGCGGGCCGTGAGTGGTCTTCAAACGAGGGTCCCGAACCCCGGGGCGGCAGGAAGGGGGCCGACT 5520

CCACGCTGTCCTTTGGGATGATACTTGGATGTCAGCTCTTGGGACCGTGTCTGCAGCCCAGCCTTCCTGTTGGGGTGGGGCCTCTCCTACTATGCAATTTTTCAAGAGCTCCTTGACCCT 5640

GCTTTTTGCTTCTTGAGTTGTCTTTTGCCATTATGGGGACTTTGGTTTGACCCAGGGGTCAGCCCTTTAGGAAGGCCTTCAGGAGGAGGCCGAGTTCCCCTTCAGTACCACCCCTCTCTC 5760

CCCACCTGCCCGCTCCCGGCAACATCTCTGGGAATCAACAGCATATTGACACGTTGGAGCCGAGCCTGAACATGCCCTGACCCCAGCACATGGAAAACCCCCTTCCTT 5868

FIG. 9A-4

1. Clone 5Y

```
G GAG TGG GGG CCA GTC ACC CGG AGC CTT CAG CGC
  E   W   G   P   V   T   R   S   L   Q   R

AGC ACC AAG CAG GAG CTG AAG GAC TTG CAG  (95)
 S   T   K   Q   E   L   K   D   L   Q  (102)
```

2. Clone 1Y

```
GGG GCC GGC GAA ACA GCG GCG GCG GCG GCG GCC CTC
 G   A   E   E   T   A   A   A   A   A   A   L

GGT GCT CTG AGG CTG GGC CGG CGG GCG CGG  (96)
 G   A   L   R   L   G   R   R   A   R  (103)
```

* - Number in parentheses is SEQ ID NO:.

FIG. 9B

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CCTACAGCTA	CCCTCACAAG	CATGAAGTGC	TGTGGCTGTT	CCTTATCCTA	50
ATGATGCGCT	TTTGTCCCGT	AAATGTTAAC	ACTCATGAAG	CATACCCCGG	100
CCTCTCAGTT	CTTGAGGGCC	TCCCCACCGC	AGCAGCAAGG	AAAGCTCACG	150
AACCCCAAAC	CTGGCAAGTC	ACCTGCAGCC	CATGGTGAGC	TCTGGGAAGT	200
GTGGTTGAGG	CCTTGGGGTC	ACTCCTTTTT	TGCATGTGCA	AATGTGCTGG	250
TCACCCTTCA	ACGCTCCCAG	ACGGTCAGGA	AAACTGTTCC	AATCATGAAA	300
AGGGGGGATG	ATTTTGTAAA	GTGGCATTTC	CTGGTCAGTG	GTGGTCTTCA	350
AGACGACAGC	TCTGTATCTG	CCATGTGAAG	AGAATTAACA	ATAAAAGTGT	400
GAAGAGCGAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	450
AAAAAAAAAA	AAAGCGCCGG	CCGC			474

FIG. 9C

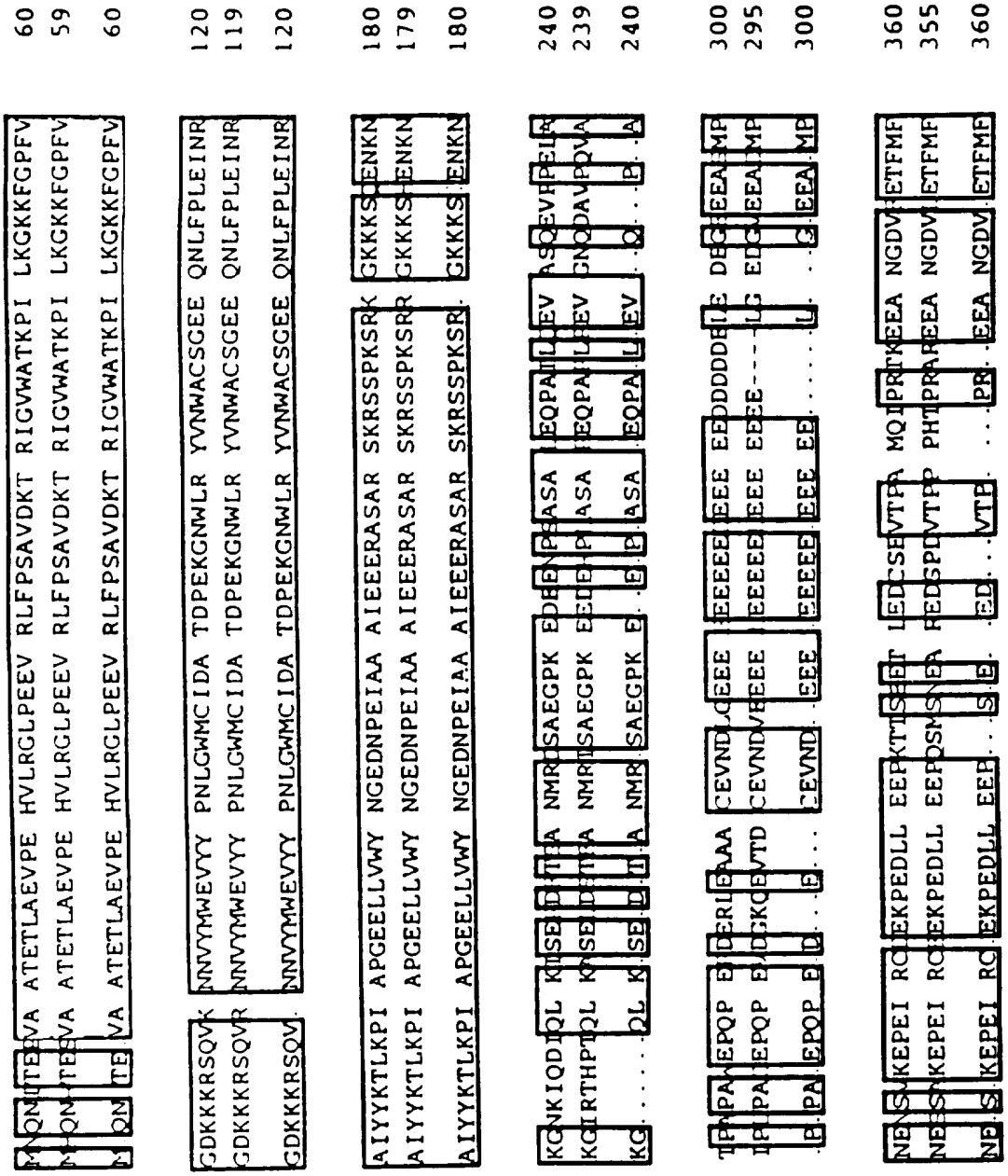

FIG. 10A

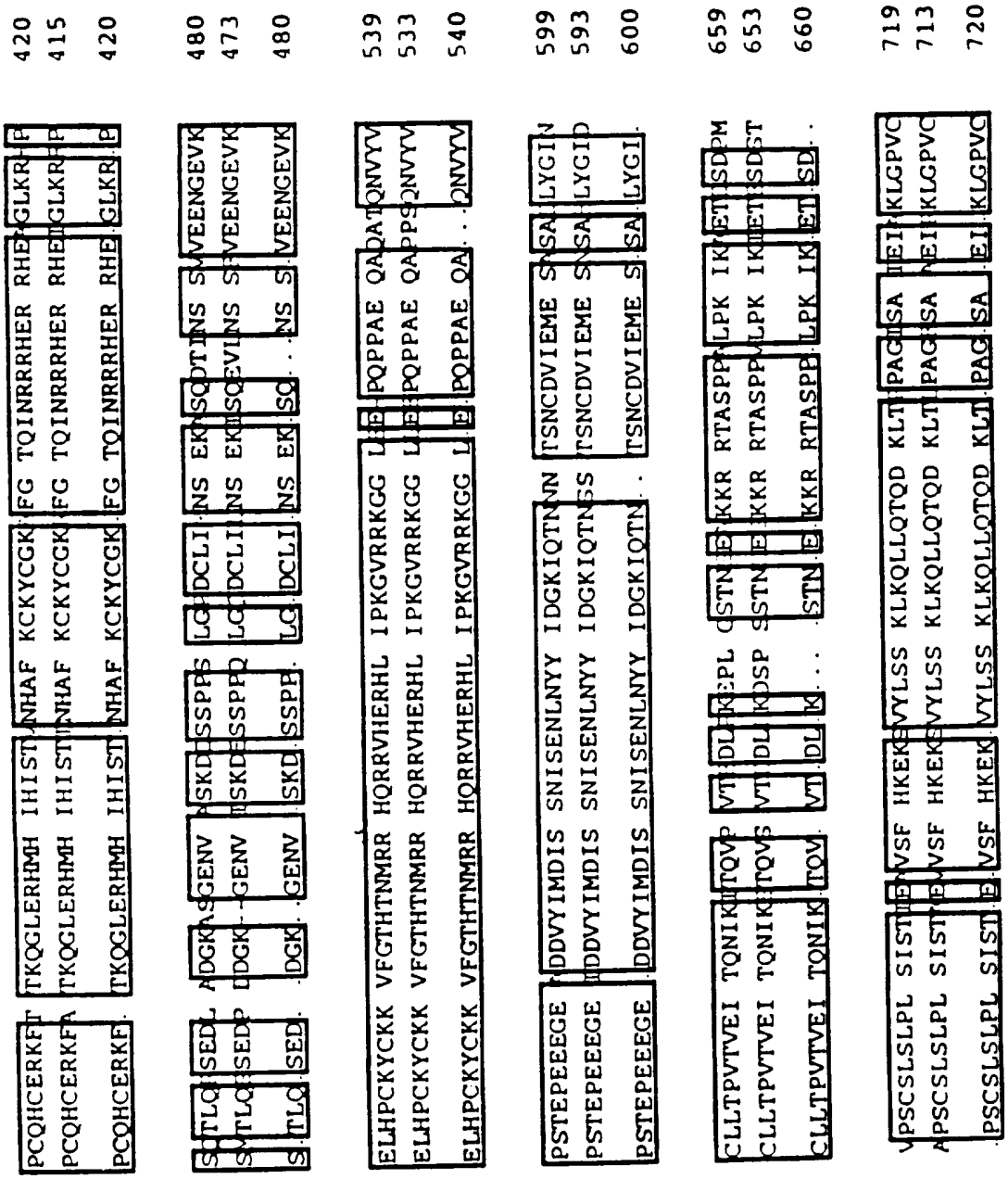

FIG. 10B

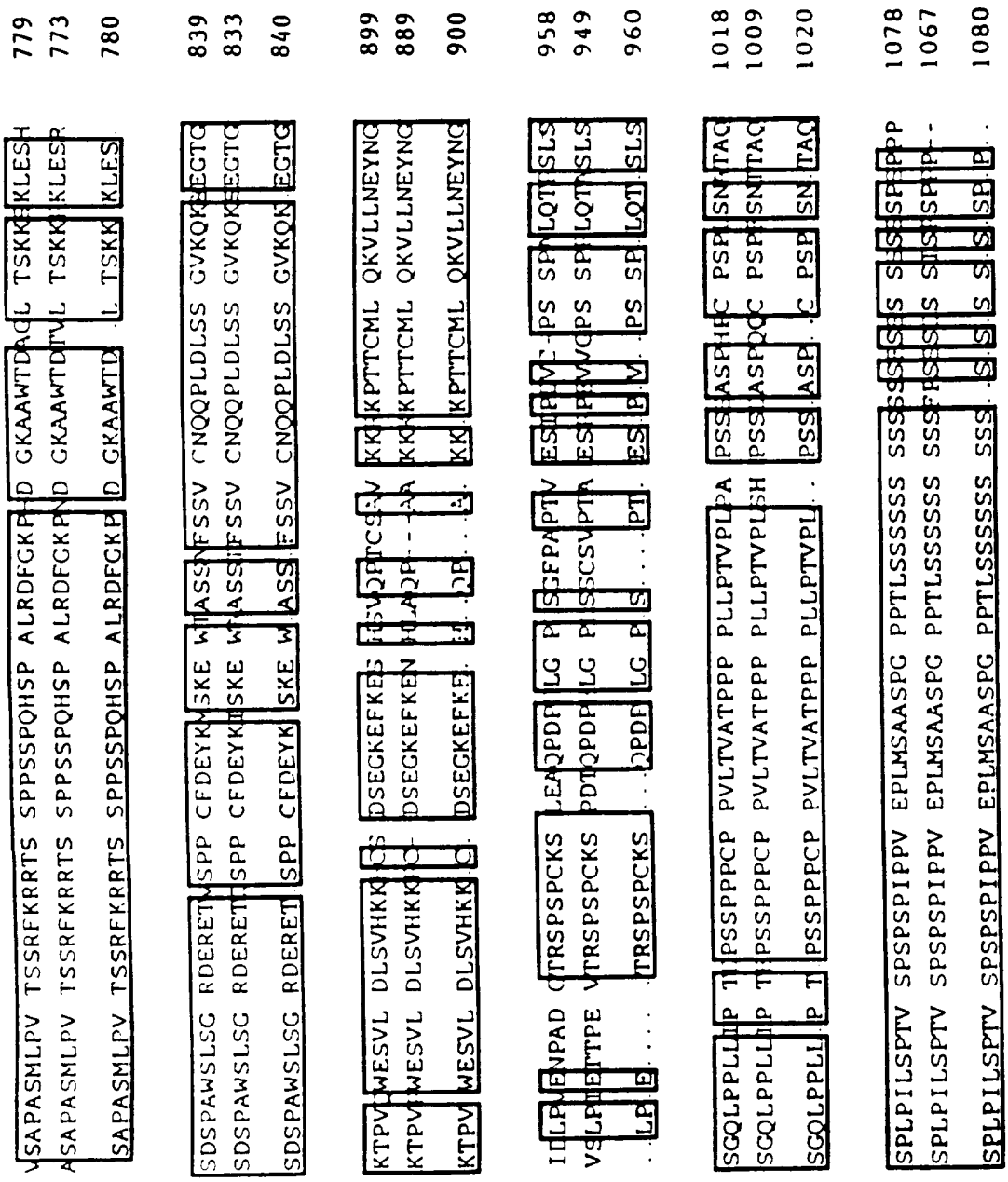
hRIZ
rRIZ
Consensus
779
773
780
839
833
840
899
889
900
958
949
960
1018
1009
1020
1078
1067
1080

FIG. 10C

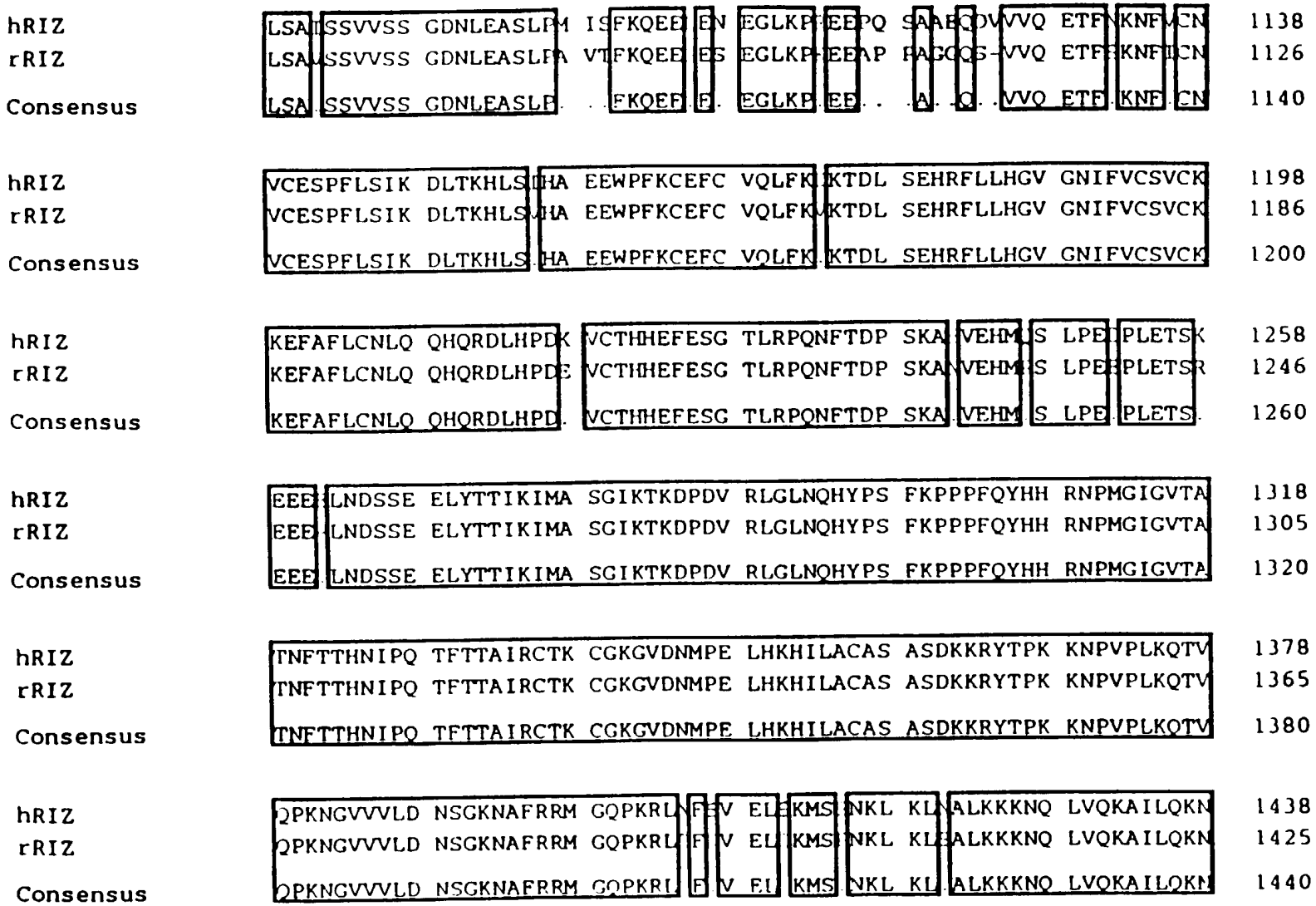

FIG. 10D

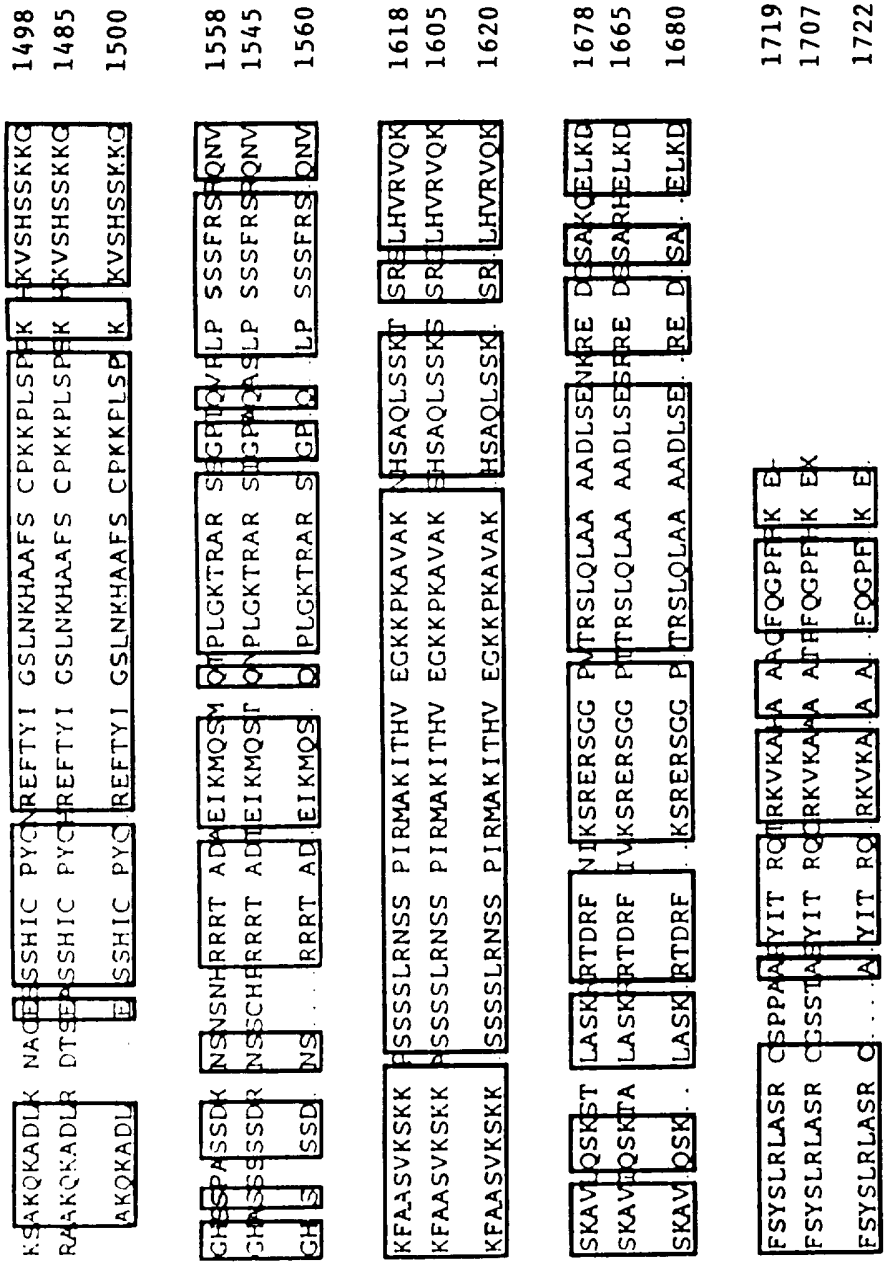
hRIZ
rRIZ
Consensus
1498
1485
1500
1558
1545
1560
1618
1605
1620
1678
1665
1680
1719
1707
1722

FIG. 10E

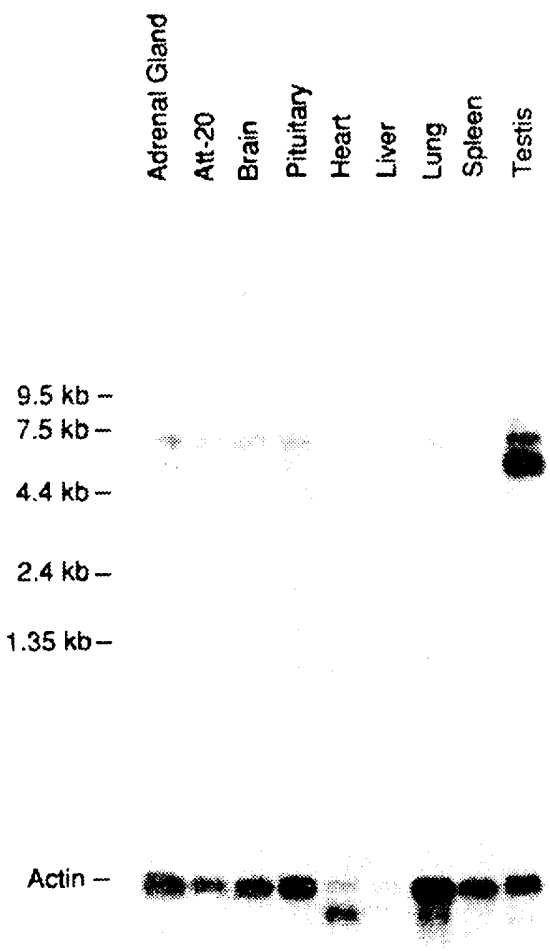

FIG. 11A

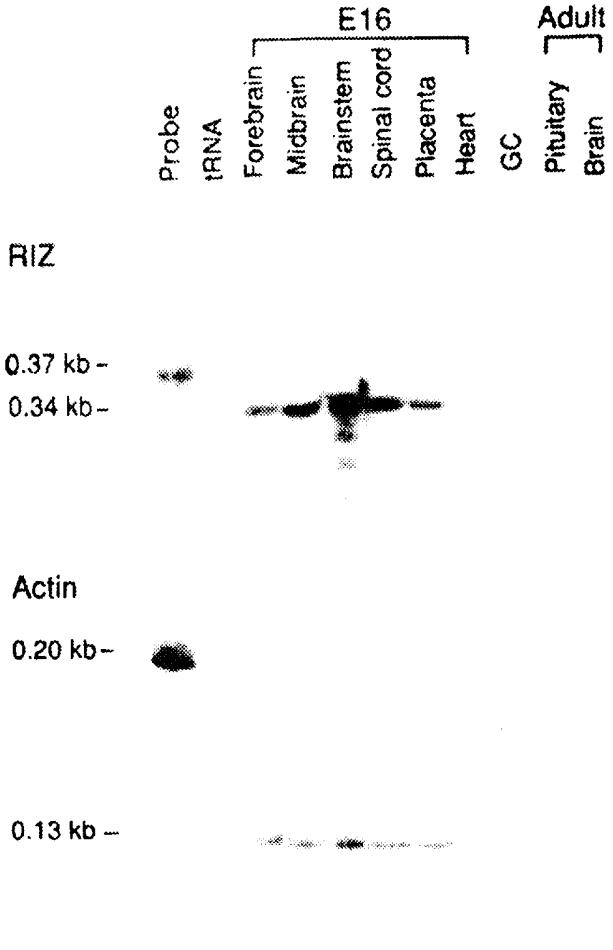

FIG. 11B

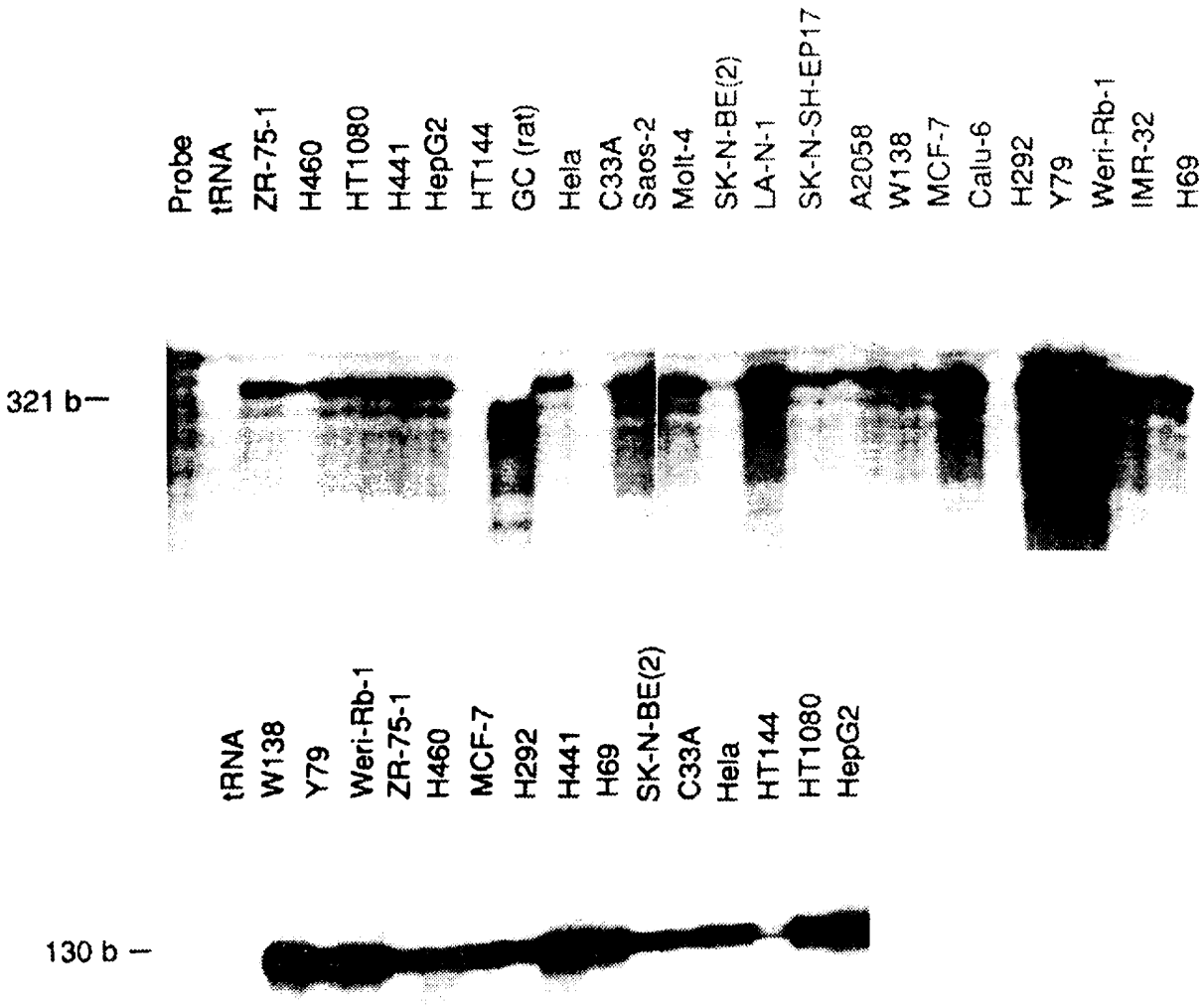
Probe
tRNA
ZR-75-1
H460
HT1080
H441
HepG2
HT144
GC (rat)
Hela
C33A
Saos-2
Molt-4
SK-N-BE(2)
LA-N-1
SK-N-SH-EP17
A2058
W138
MCF-7
Calu-6
H292
Y79
Weri-Rb-1
IMR-32
H69
321 b
tRNA
W138
Y79
Weri-Rb-1
ZR-75-1
H460
MCF-7
H292
H441
H69
SK-N-BE(2)
C33A
Hela
HT144
HT1080
HepG2
130 b

FIG. 12

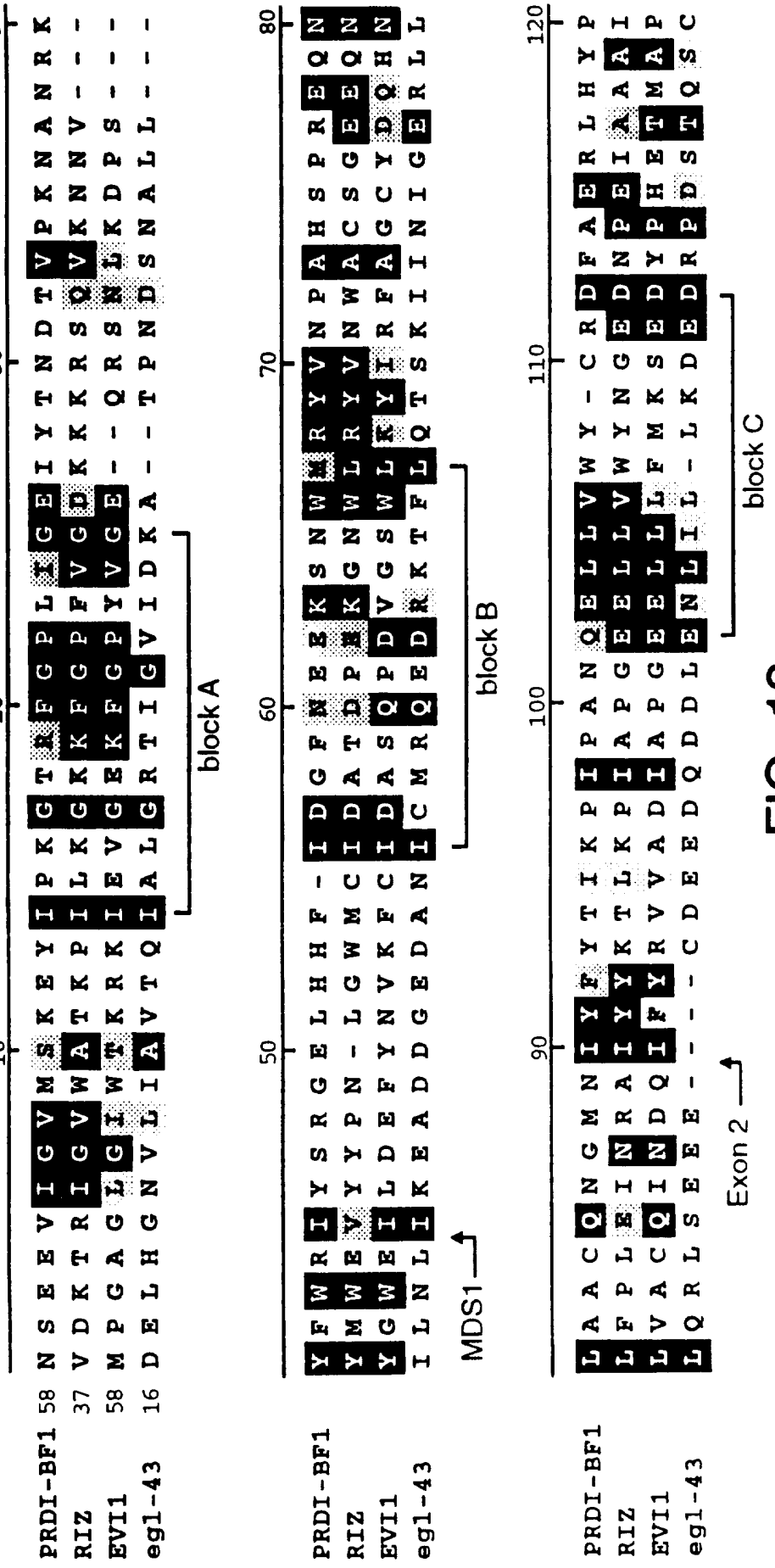
PRDI-BF1 58 N S E E V I G V M S K E Y I P K G T R F G P L I G E I Y T N D T V P K N A N R K
RIZ 37 V D K T R I G V W A T K P I L K G K F G P F V G D K K K R S Q V K N N V - - - -
EVI1 58 M P G A G L G I W T K R K I E V G K F G P Y V G E - - Q R S N L K D P S - - - -
egl-43 16 D E L H G N V L I A V T Q I A L G R T I G V I D K A - - T P N D S N A L L - - -
block A
10
20
30
40
PRDI-BF1 Y F W R I Y S R G E L H H F - I D G F N E E K S N W M R Y V N P A H S P R E Q N
RIZ Y M W E V Y Y P N - L G W M C I D A T D P E K G N W L R Y V N W A C S G E E Q N
EVI1 Y G W E I L D E F Y N V K F C I D A S Q P D V G S W L K Y I R F A G C Y D Q H N
egl-43 I L N L I K E A D D G E D A N I C M R Q E D R K T F L Q T S K I I N I G E R L L
MDS1
block B
50
60
70
80
PRDI-BF1 L A A C Q N G M N I Y F Y T I K P I P A N Q E L L V W Y - C R D F A E R L H Y P
RIZ L F P L E I N R A I Y Y K T L K P I A P G E E L L V W Y N G E D N P E I A A A I
EVI1 L V A C Q I N D Q I F Y R V V A D I A P G E E L L F M K S E D Y P H E T M A P
egl-43 L Q R L S E E E - - - - C D E E D Q D D L E N L I L - L K D E D R P D S T Q S C
Exon 2
block C
90
100
110
120

FIG. 13

```
          10        20        30        40        50        60
1234567890123456789012345678901234567890123456789012345678901234567890

GTGTACTACCCAAATTTGGGGTGGATGTGCATTGATGCCACTGATCCGGAGAAGGGCAAC                60
ValTyrTyrProAsnLeuGlyTrpMetCysIleAspAlaThrAspProGluLysGlyAsn

TGGCTCCGCTATGTGAACTGGGCTTGCTCAGGAGAAGAACAGAATTTATTTCCACTGGAA               120
TrpLeuArgTyrValAsnTrpAlaCysSerGlyGluGluGlnAsnLeuPheProLeuGlu

ATCAACAGAGCCATTTACTATAAAACCTTAAAGCCAATCGCGCCTGGCGAGGAGCTCCTG               180
IleAsnArgAlaIleTyrTyrLysThrLeuLysProIleAlaProGlyGluGluLeuLeu

GTCTGGTACAATGGGGAAGACAACCCCGAGATAGCAGCTGCGATTGAGGAAGAGCGAGCC               240
ValTrpTyrAsnGlyGluAspAsnProGluIleAlaAlaAlaIleGluGluGluArgAla

AGCGCCCGGAGCAAGCGGAGCTCCCCGAAGAGCCGGAGAGGGAAGAAGAAATCACAGGAG               300
SerAlaArgSerLysArgSerSerProLysSerArgArgGlyLysLysLysSerGlnGlu

AATAAAAACAAAGGCATCAGAACCCAGGCTGCAGCGCGGAAGGCGAGCGAGCTGGACTCC               360
AsnLysAsnLysGluIleArgThrGlnAlaAlaAlaArgLysAlaSerGluLeuAspSer

ACCTCTGCAAACATGAGGGGCTCTGCAGAAG                                              391
ThrSerAlaAsnMetArgGlySerAlaGlu
```

FIG. 14