



US 20080131920A1

(19) United States

(12) Patent Application Publication
Levin

(10) Pub. No.: US 2008/0131920 A1

(43) Pub. Date: Jun. 5, 2008

(54) ION FLUX IN BIOLOGICAL PROCESSES,
AND METHODS RELATED THERETO(75) Inventor: Michael Levin, Swampscott, MA
(US)

Correspondence Address:
ROPES & GRAY LLP
PATENT DOCKETING 39/41, ONE INTERNAL
PLACE
BOSTON, MA 02110-2624

(73) Assignee: The Forsyth Institute, Boston, MA
(US)

(21) Appl. No.: 11/543,469

(22) Filed: Oct. 4, 2006

Related U.S. Application Data

(60) Provisional application No. 60/723,414, filed on Oct. 4, 2005, provisional application No. 60/841,777, filed on Aug. 31, 2006.

Publication Classification

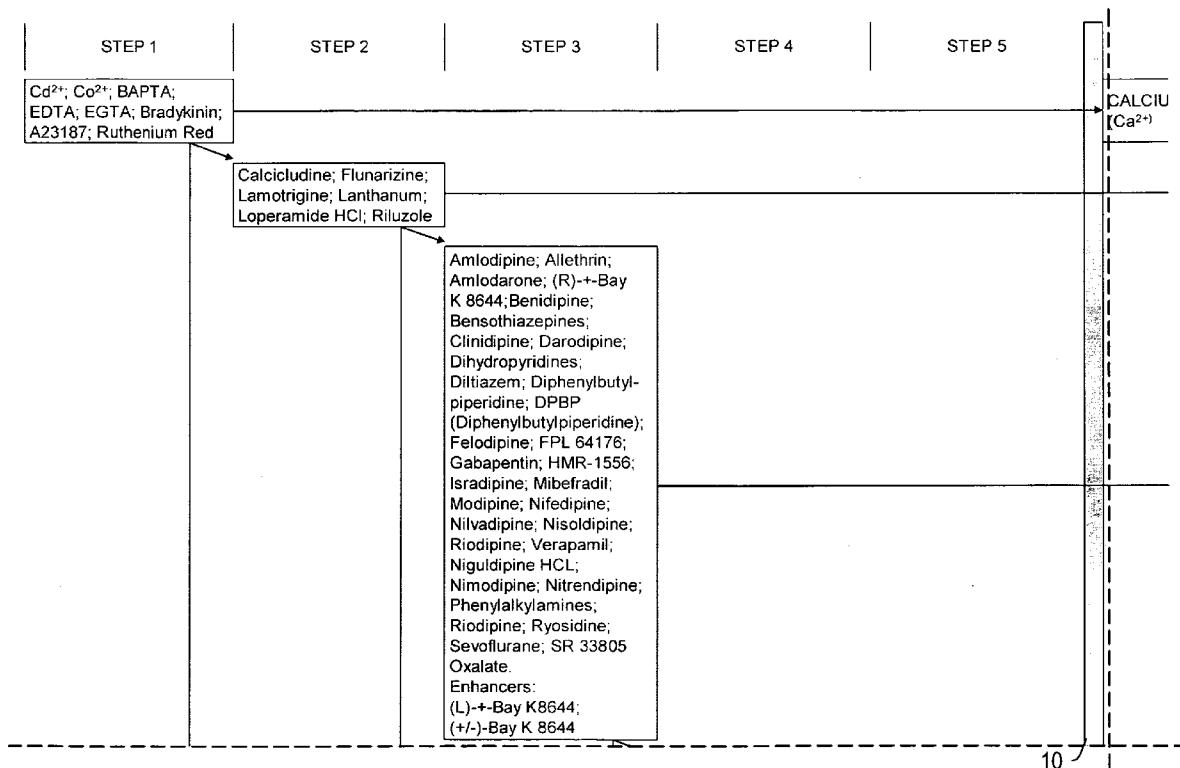
(51) Int. Cl.

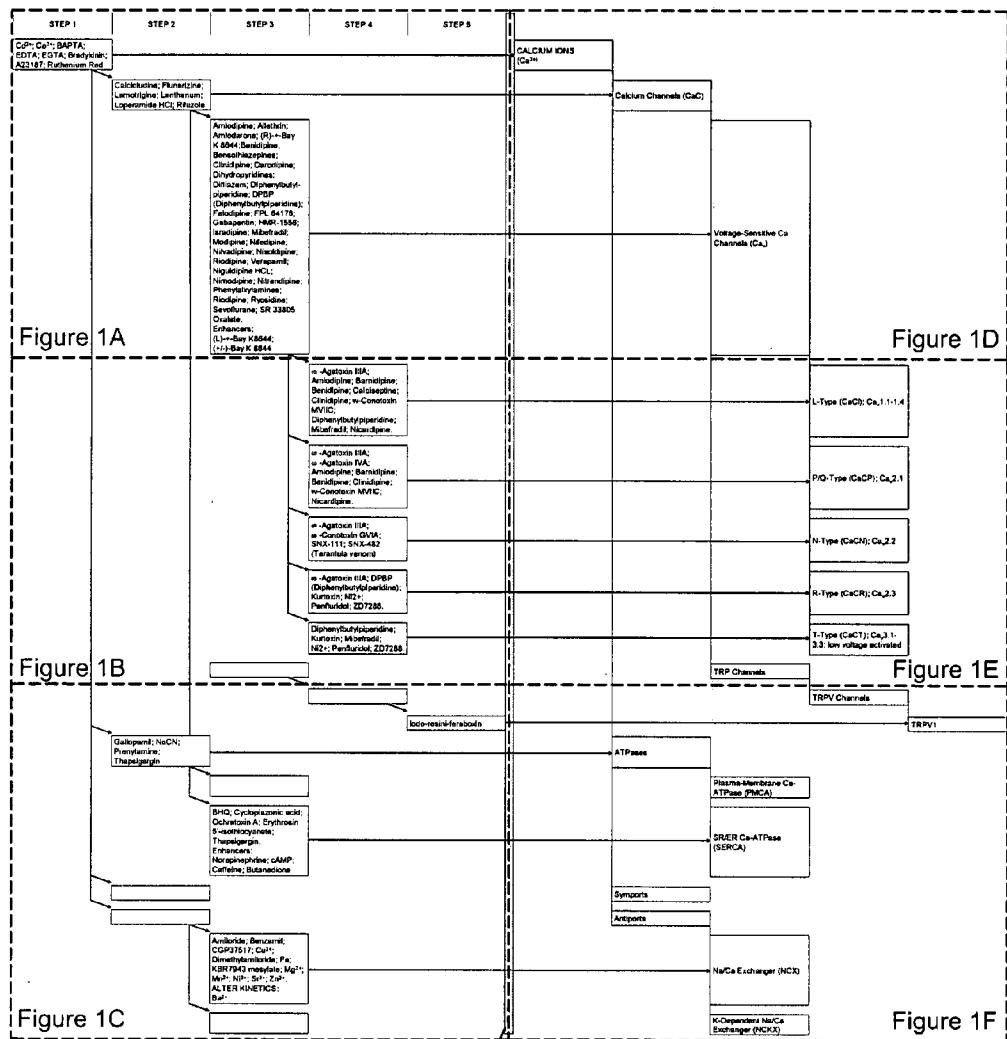
C12N 5/00 (2006.01)
C12N 5/02 (2006.01)
C12Q 1/20 (2006.01)
C12N 5/04 (2006.01)

(52) U.S. Cl. 435/29; 435/375; 435/377; 435/378

(57) ABSTRACT

The present invention provides methods for promoting differentiation and/or regeneration by modulating membrane potential and/or intracellular pH in non-naturally regenerating cells.





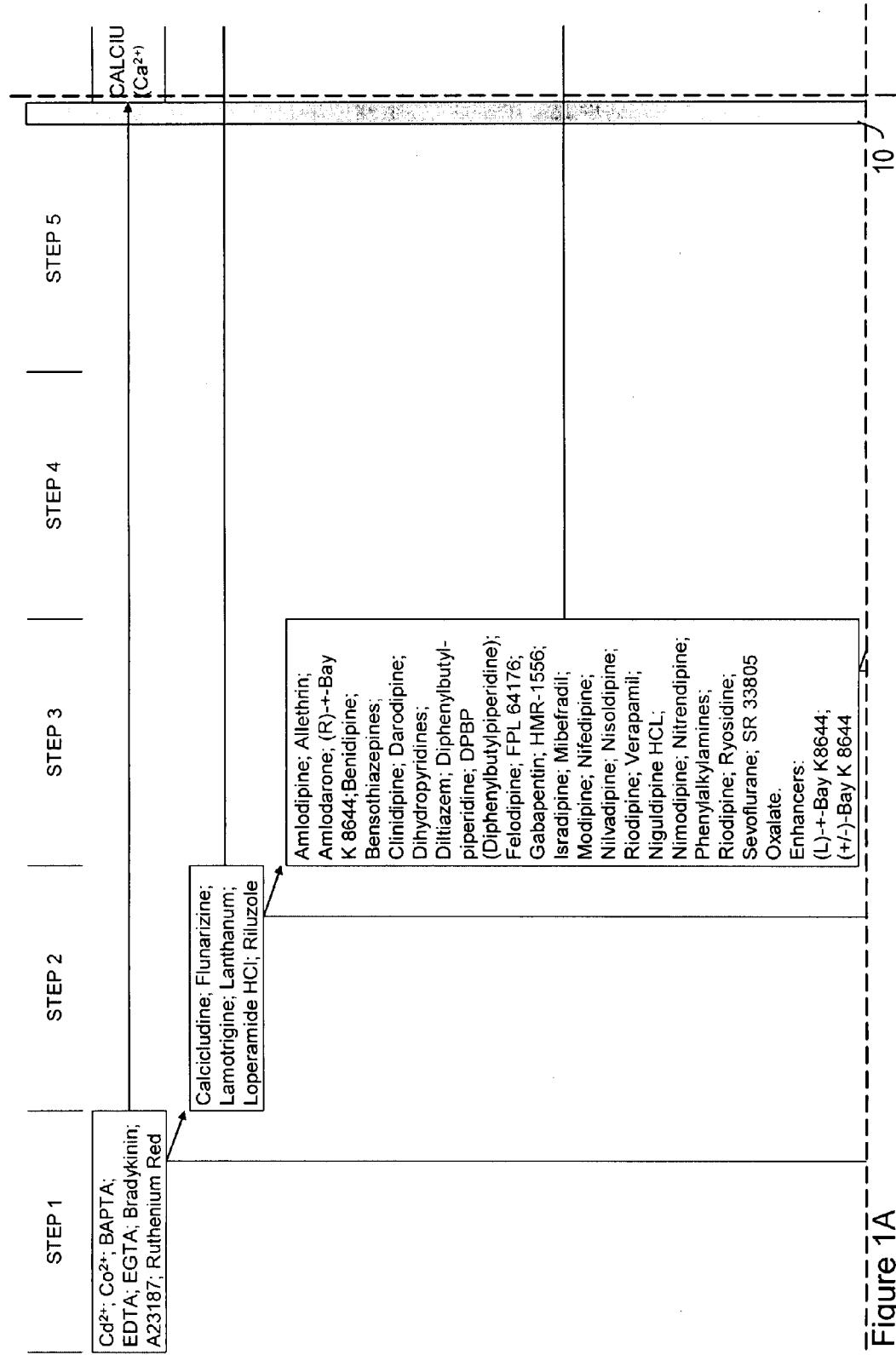


Figure 1A

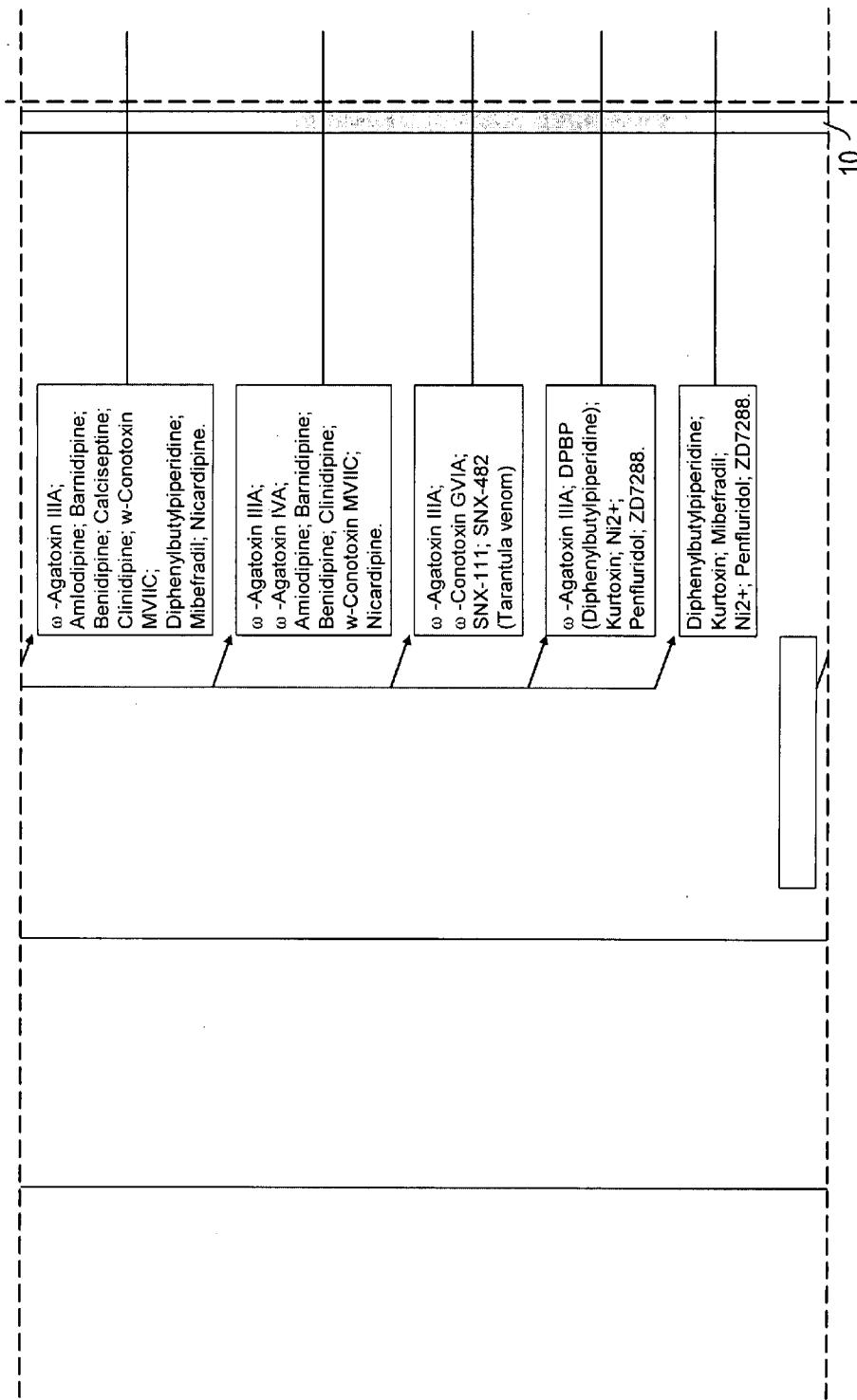


Figure 1B

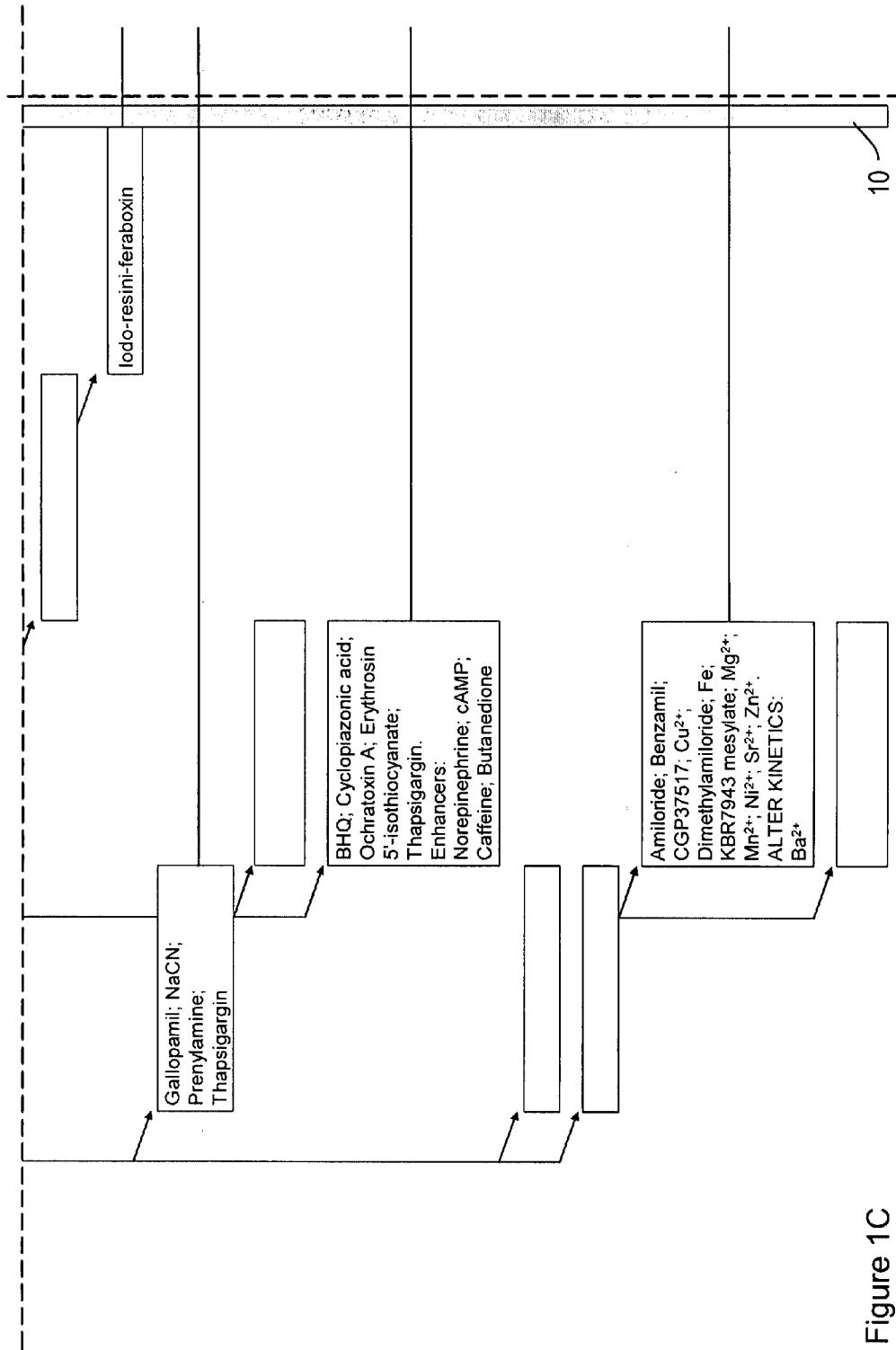
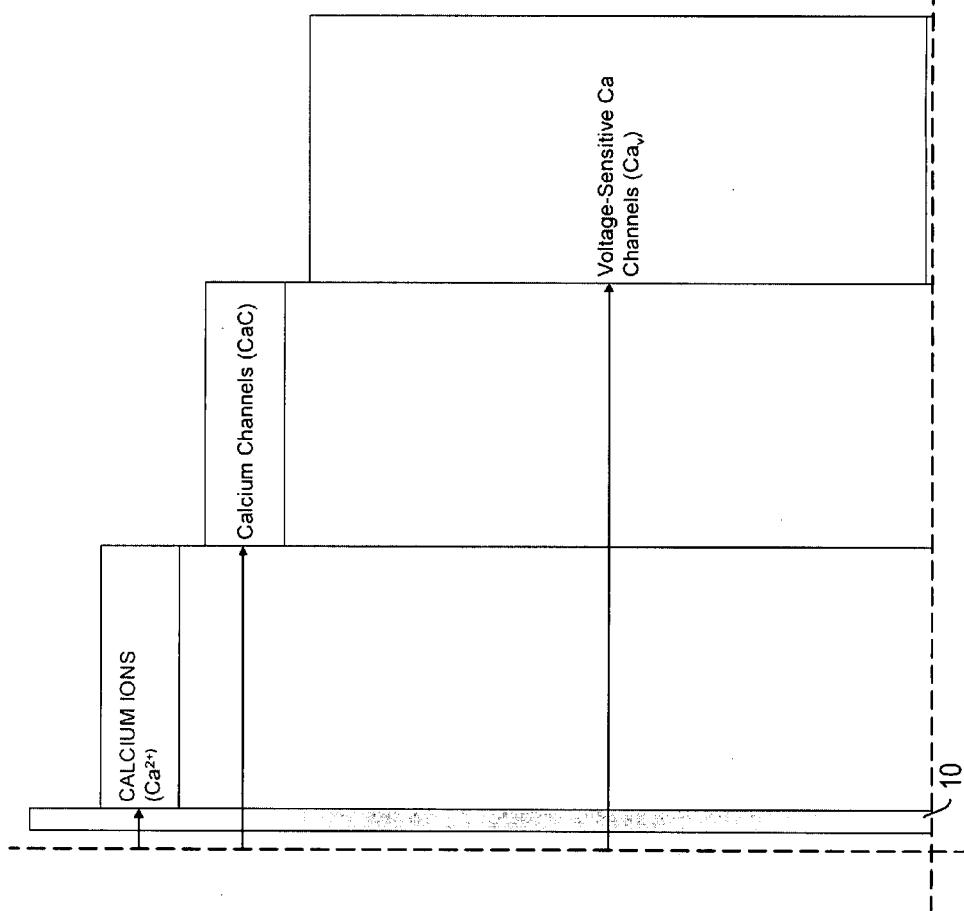


Figure 1C

Figure 1D



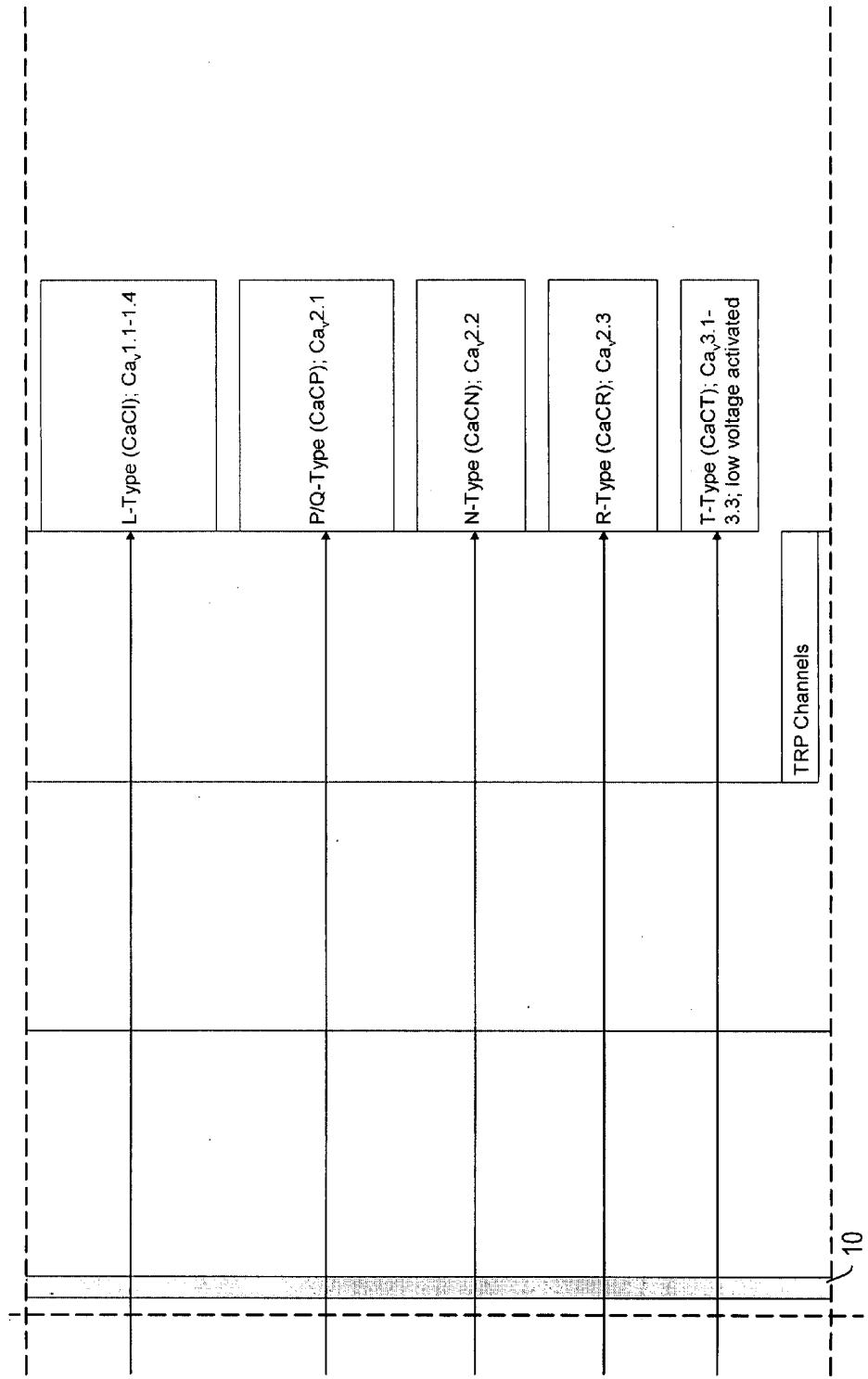


Figure 1E

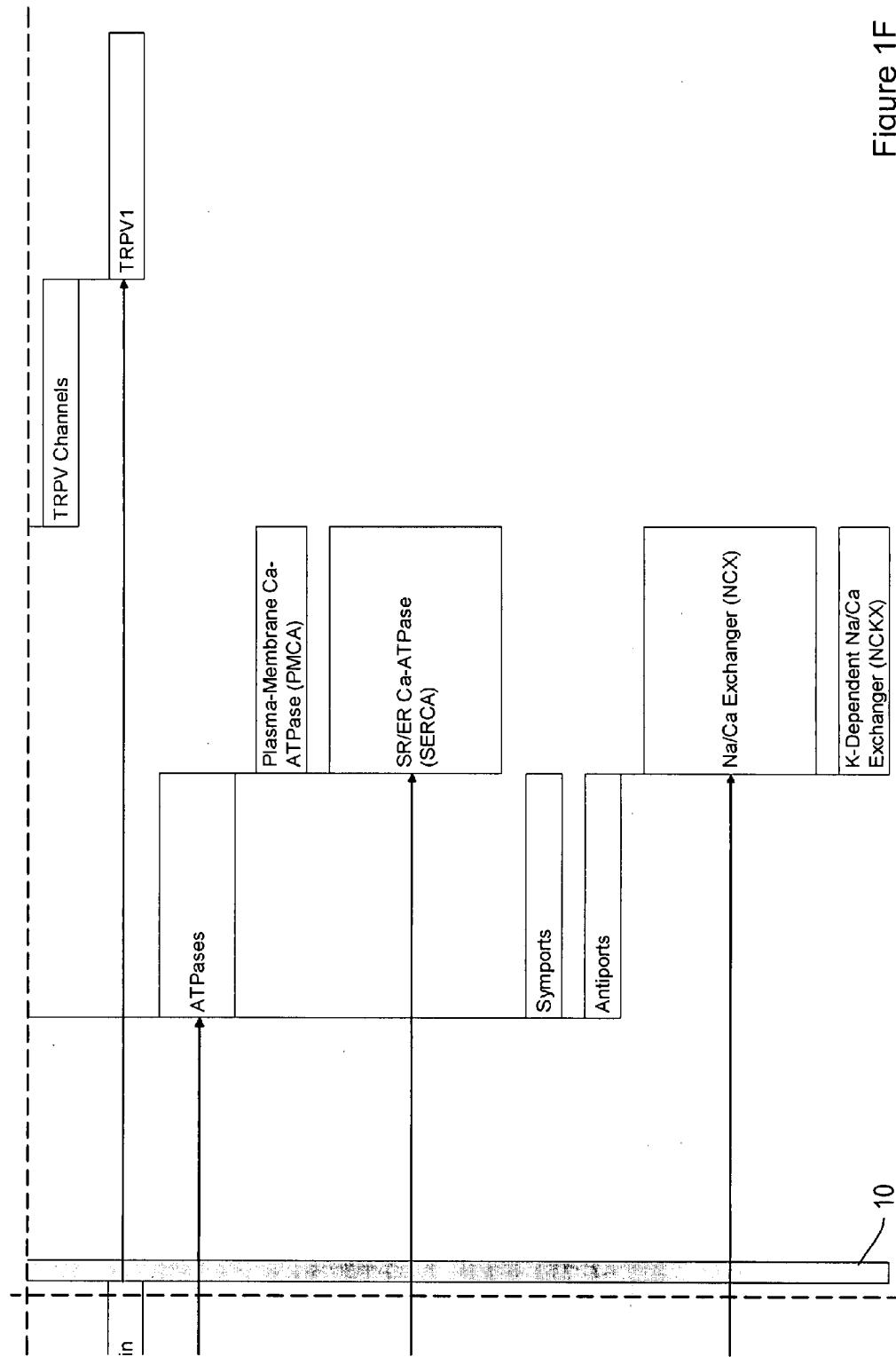
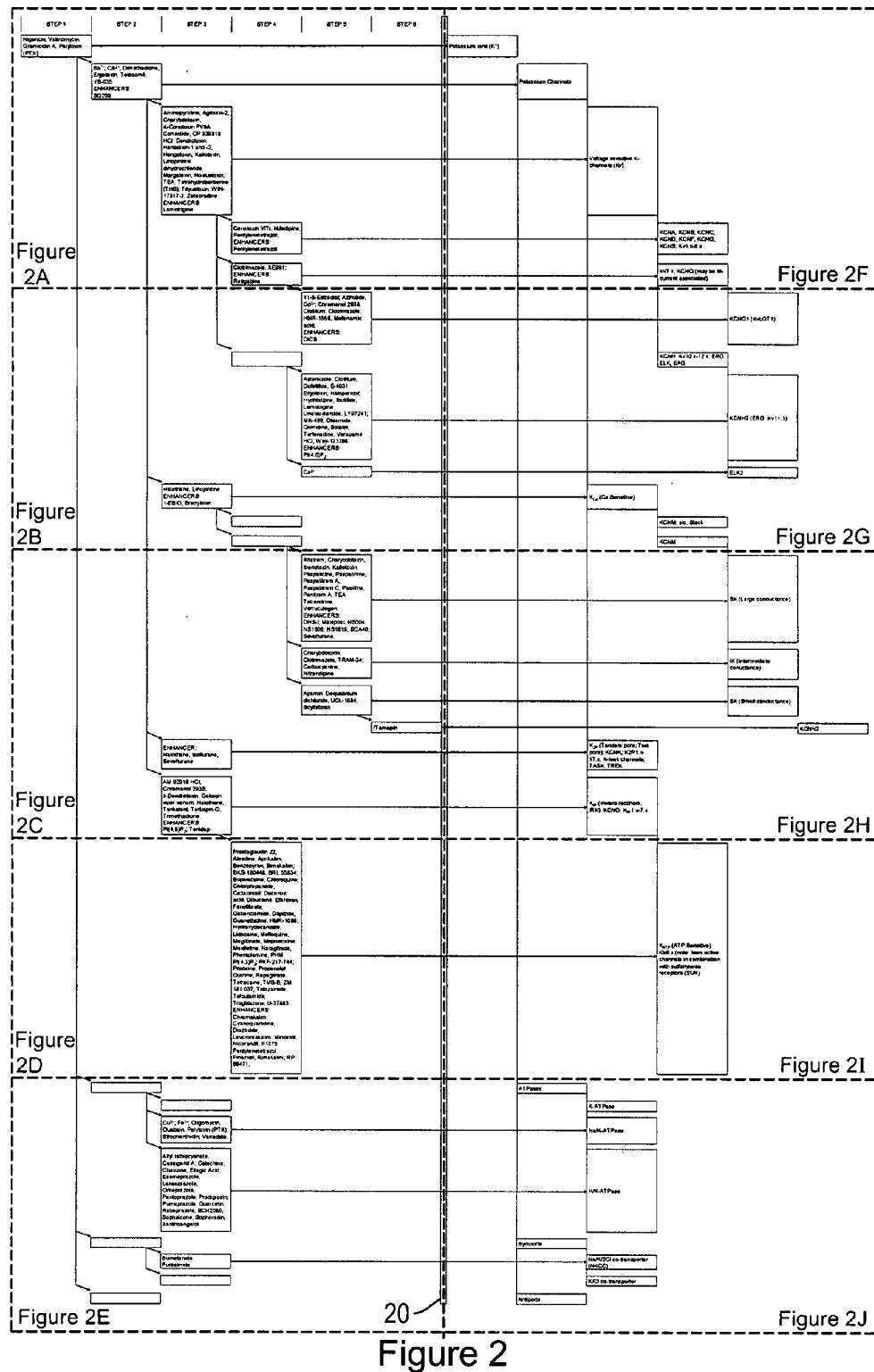


Figure 1F



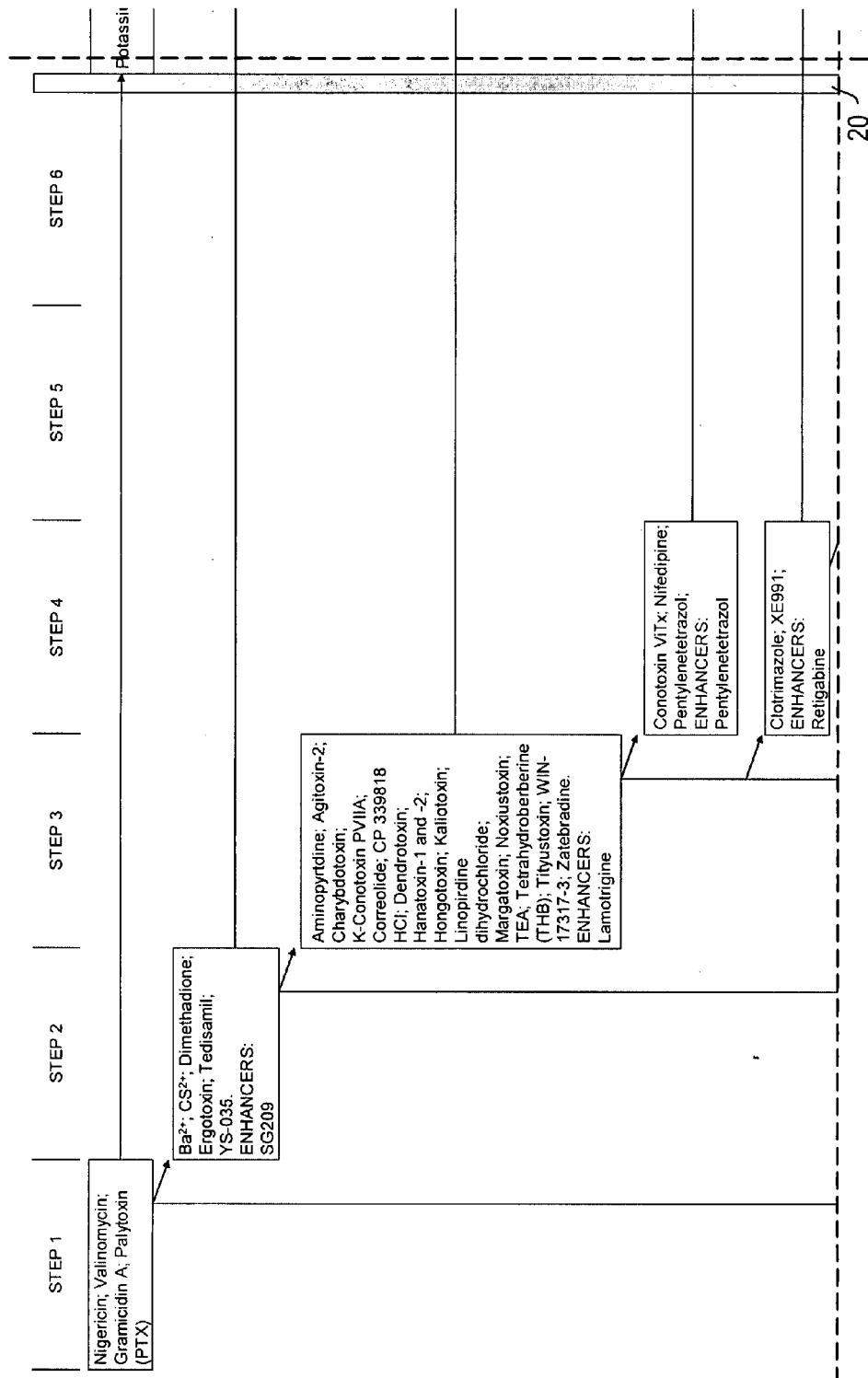


Figure 2A

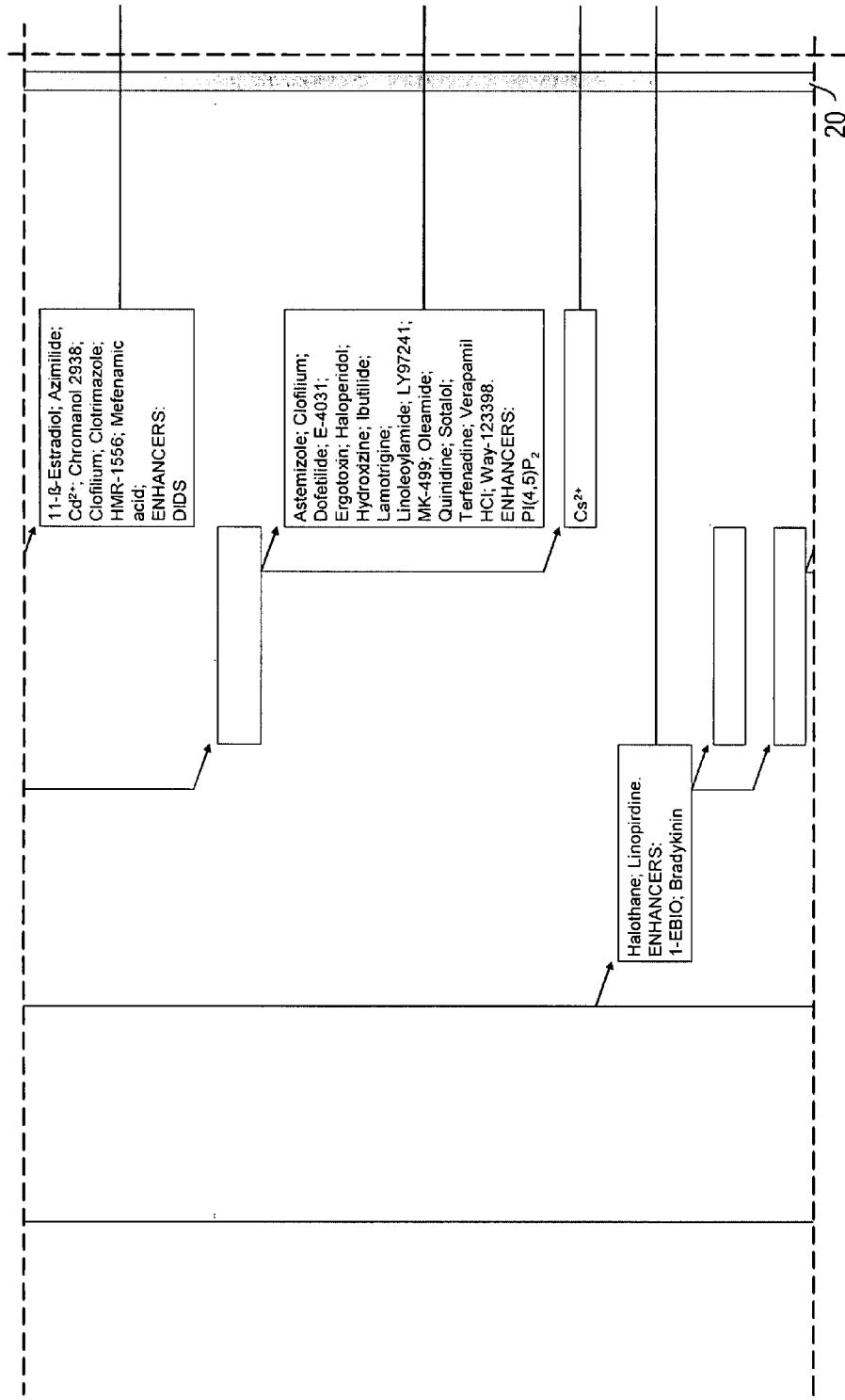


Figure 2B

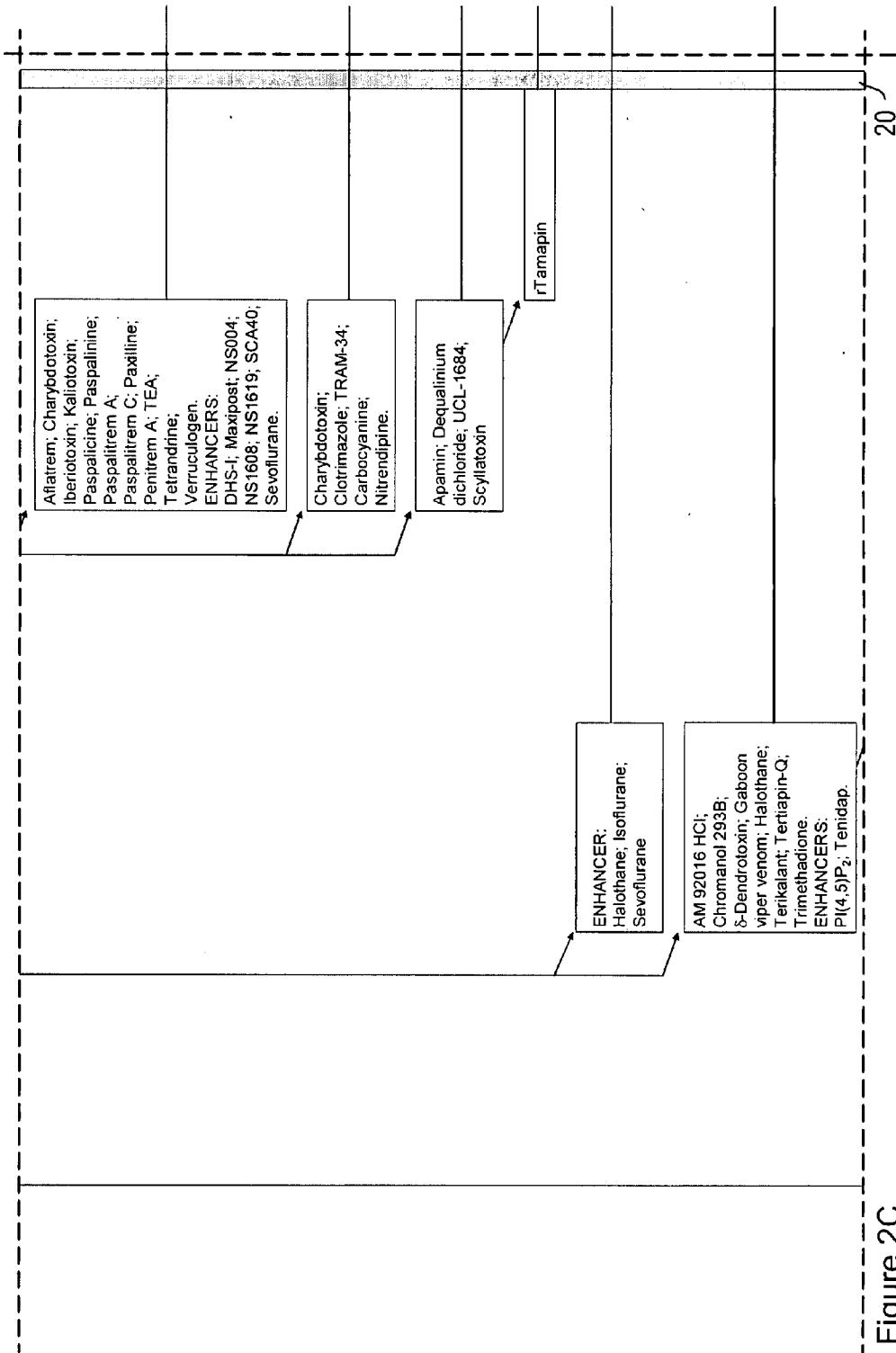


Figure 2C

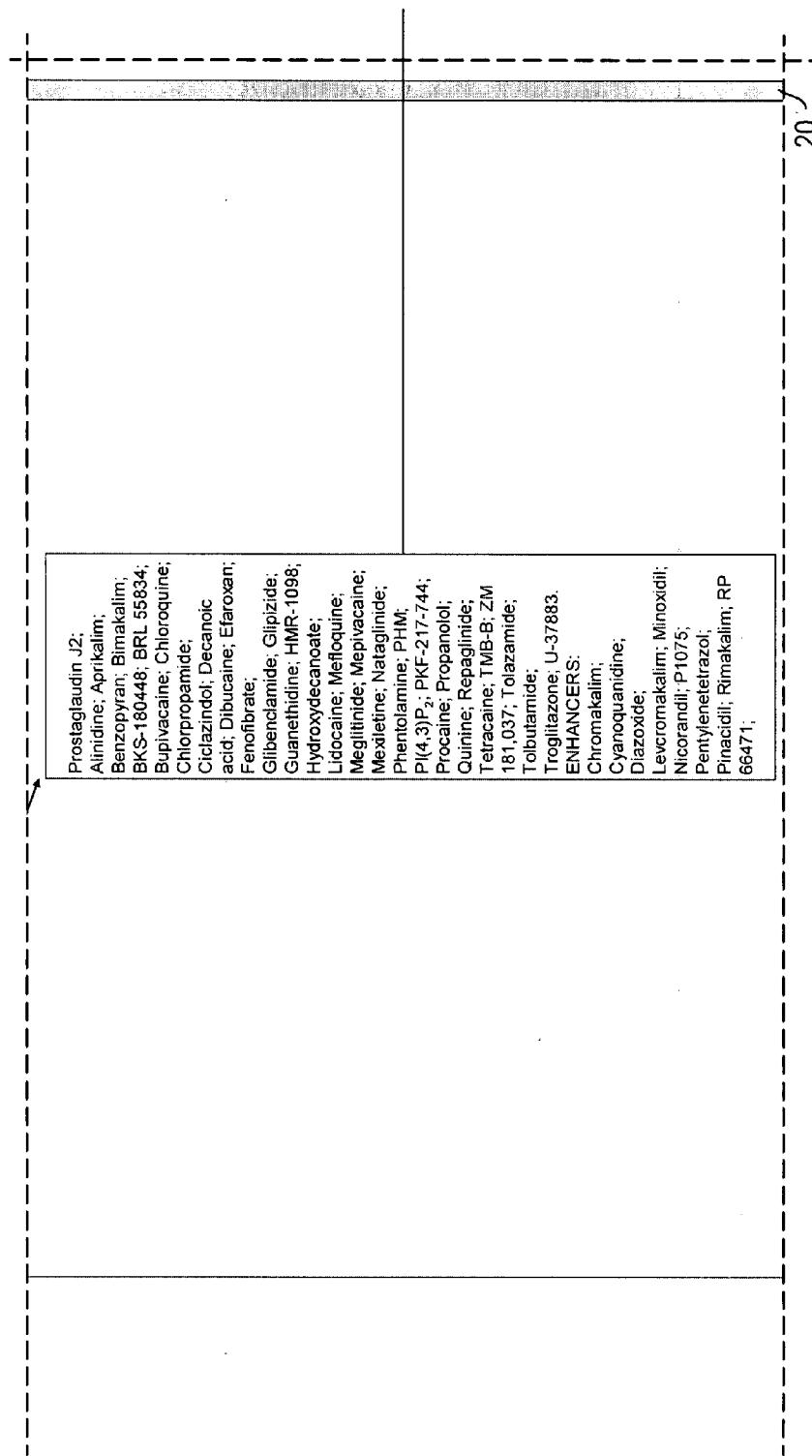


Figure 2D

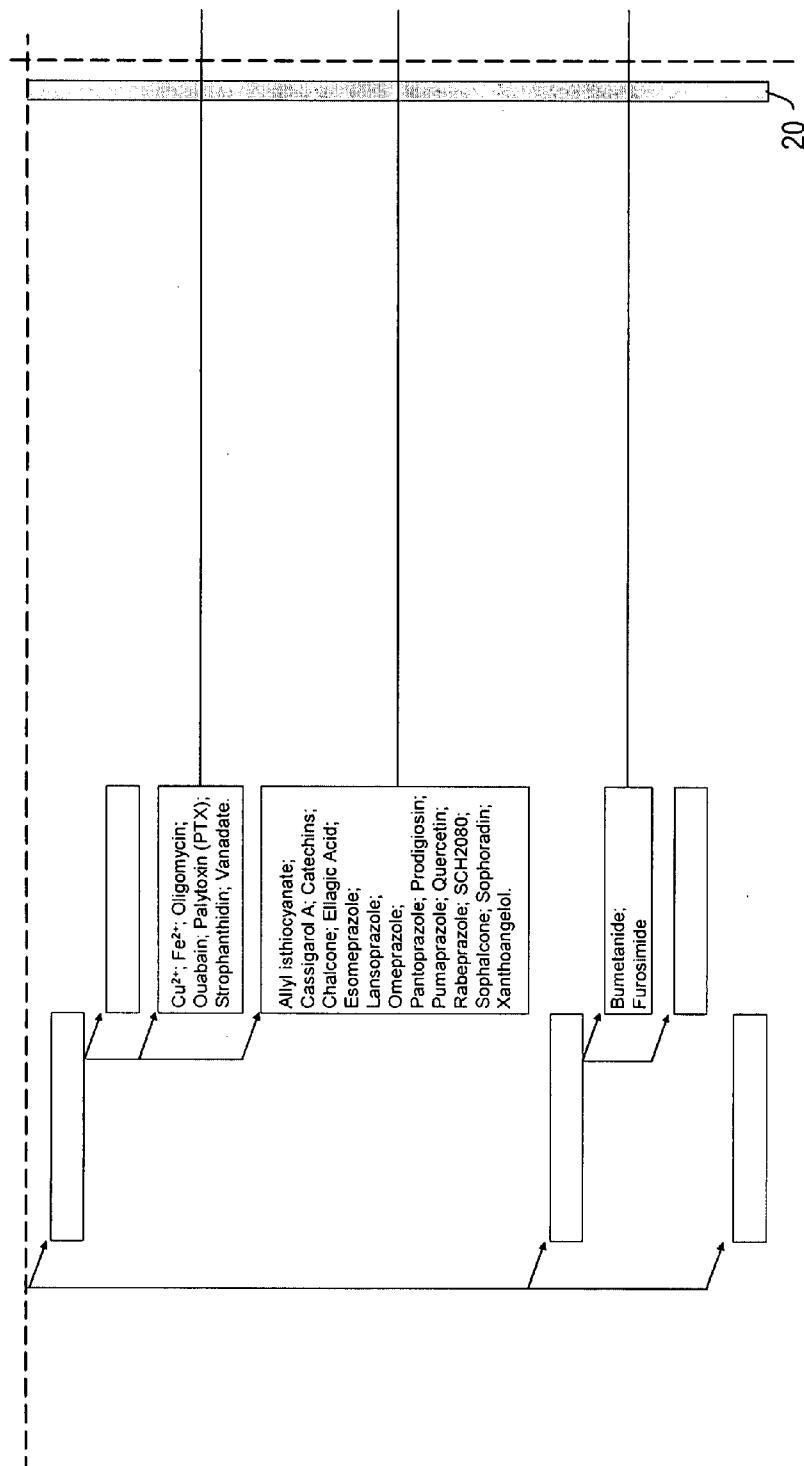


Figure 2E

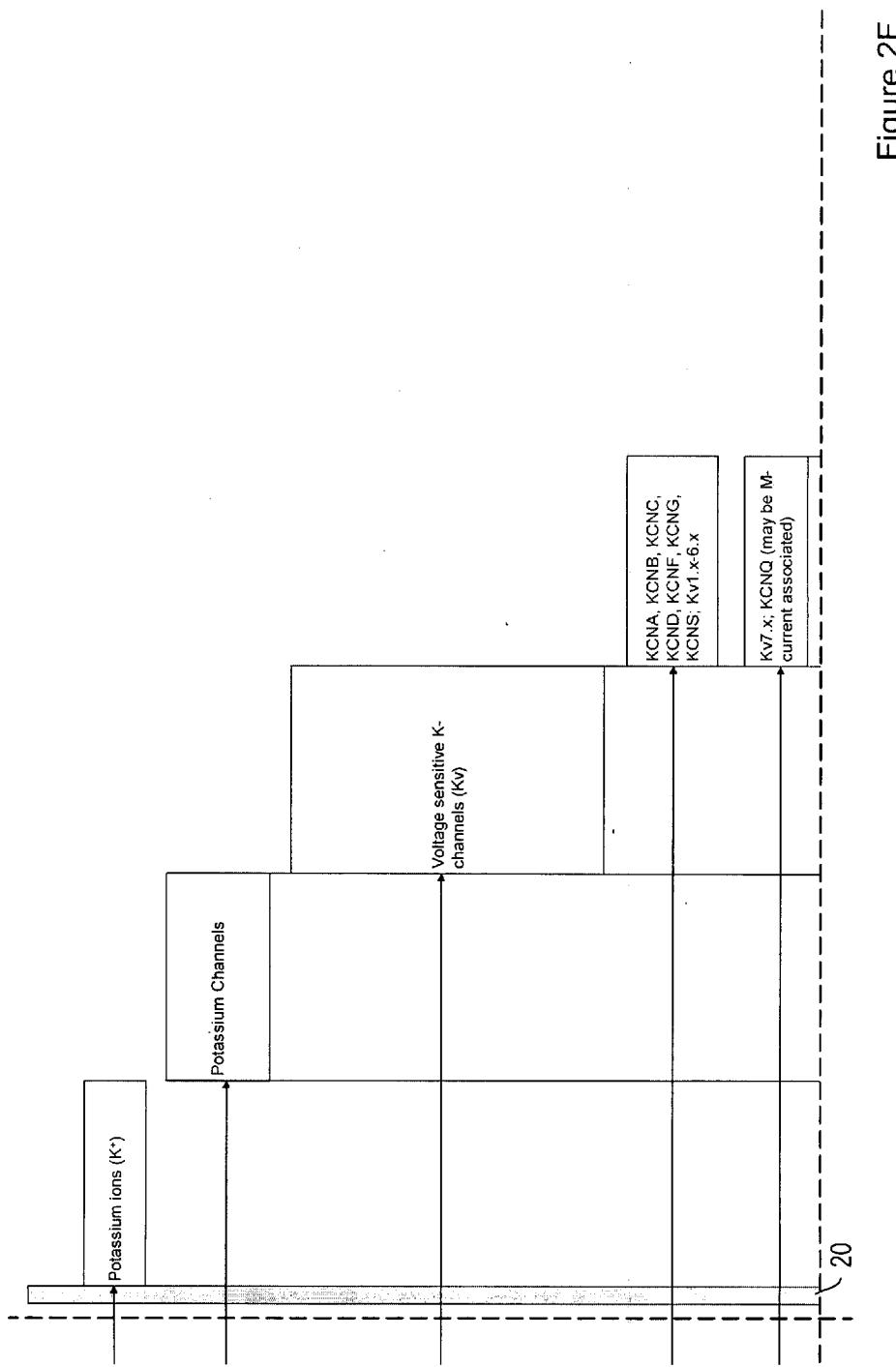


Figure 2F

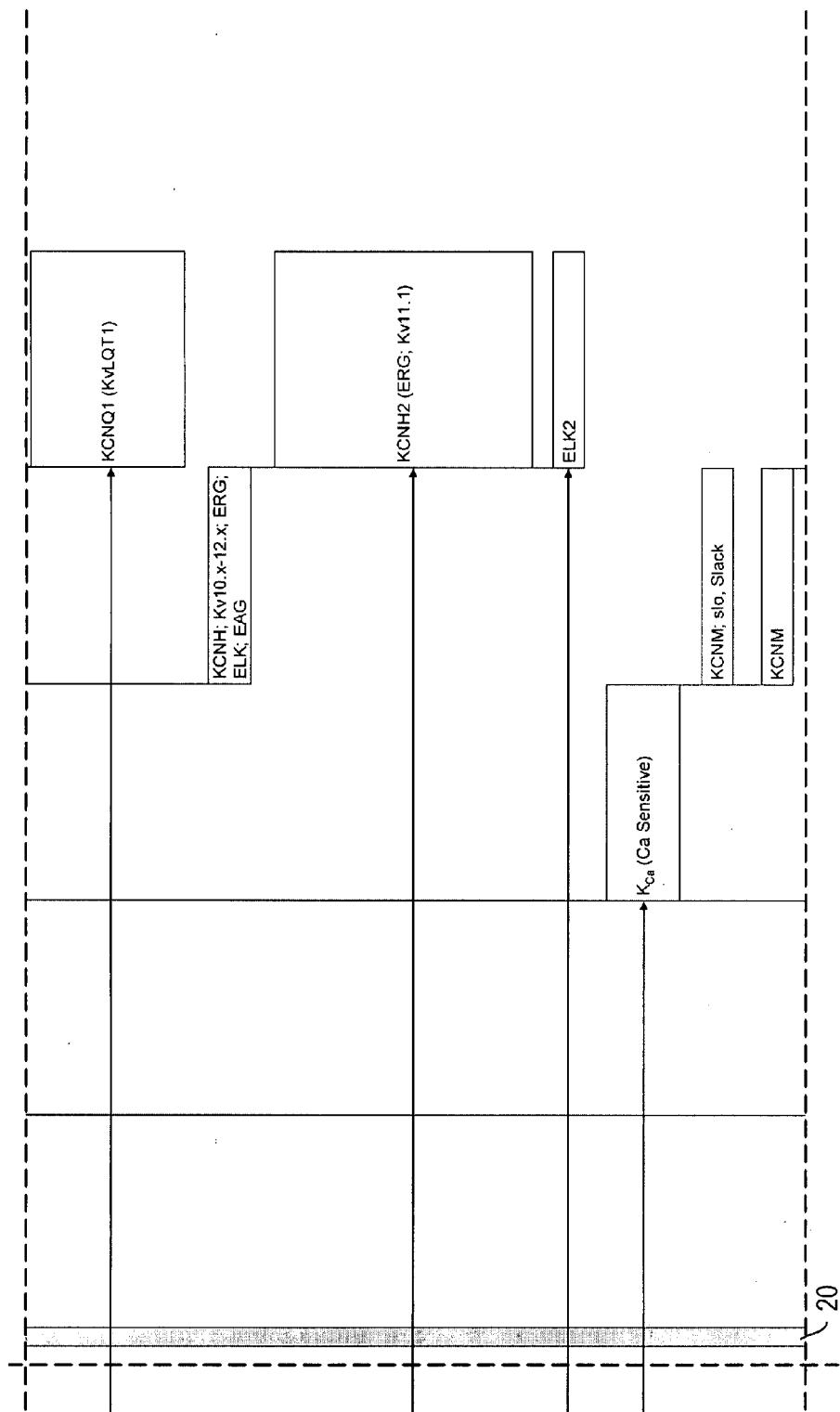


Figure 2G

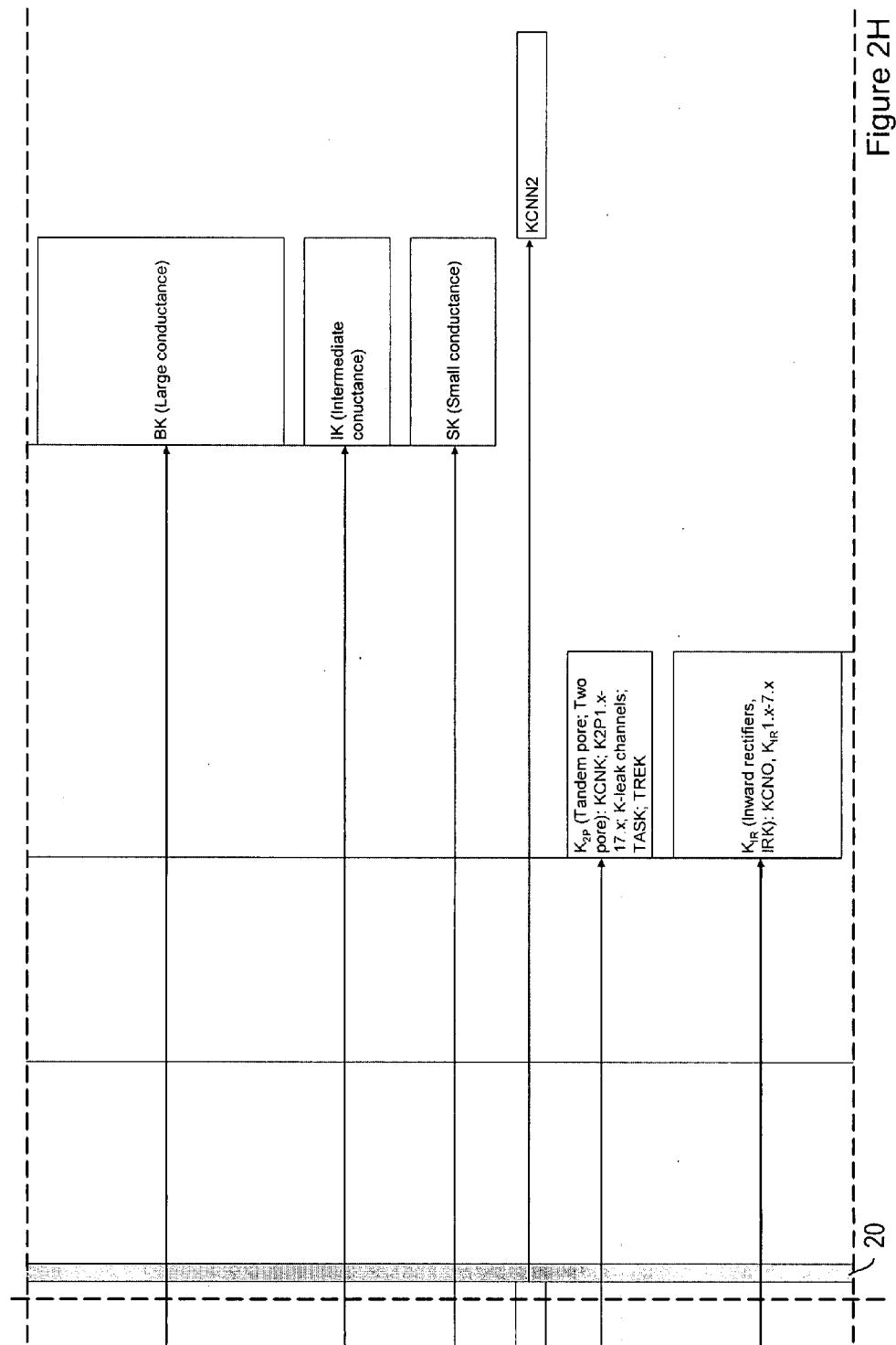


Figure 2H

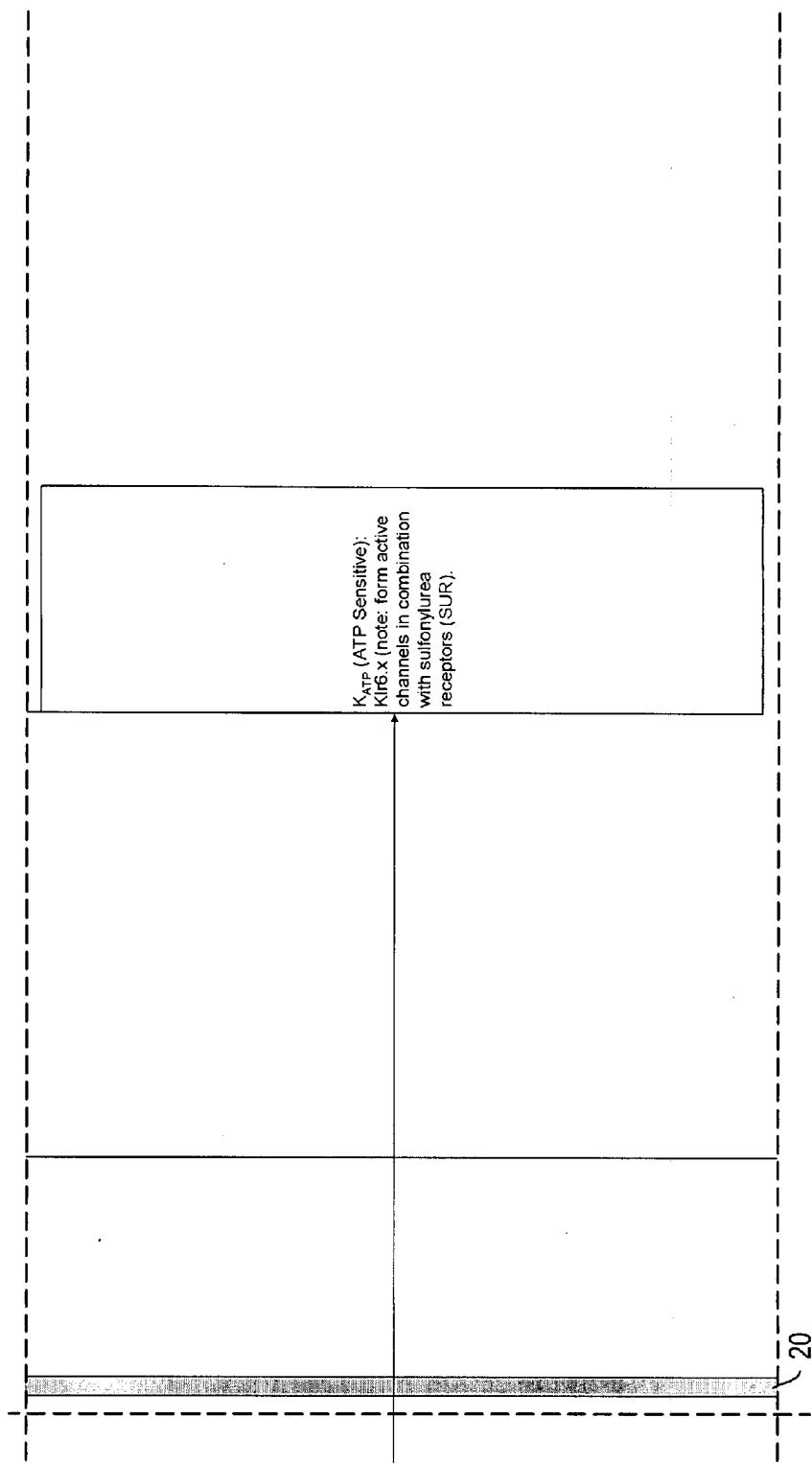


Figure 2I

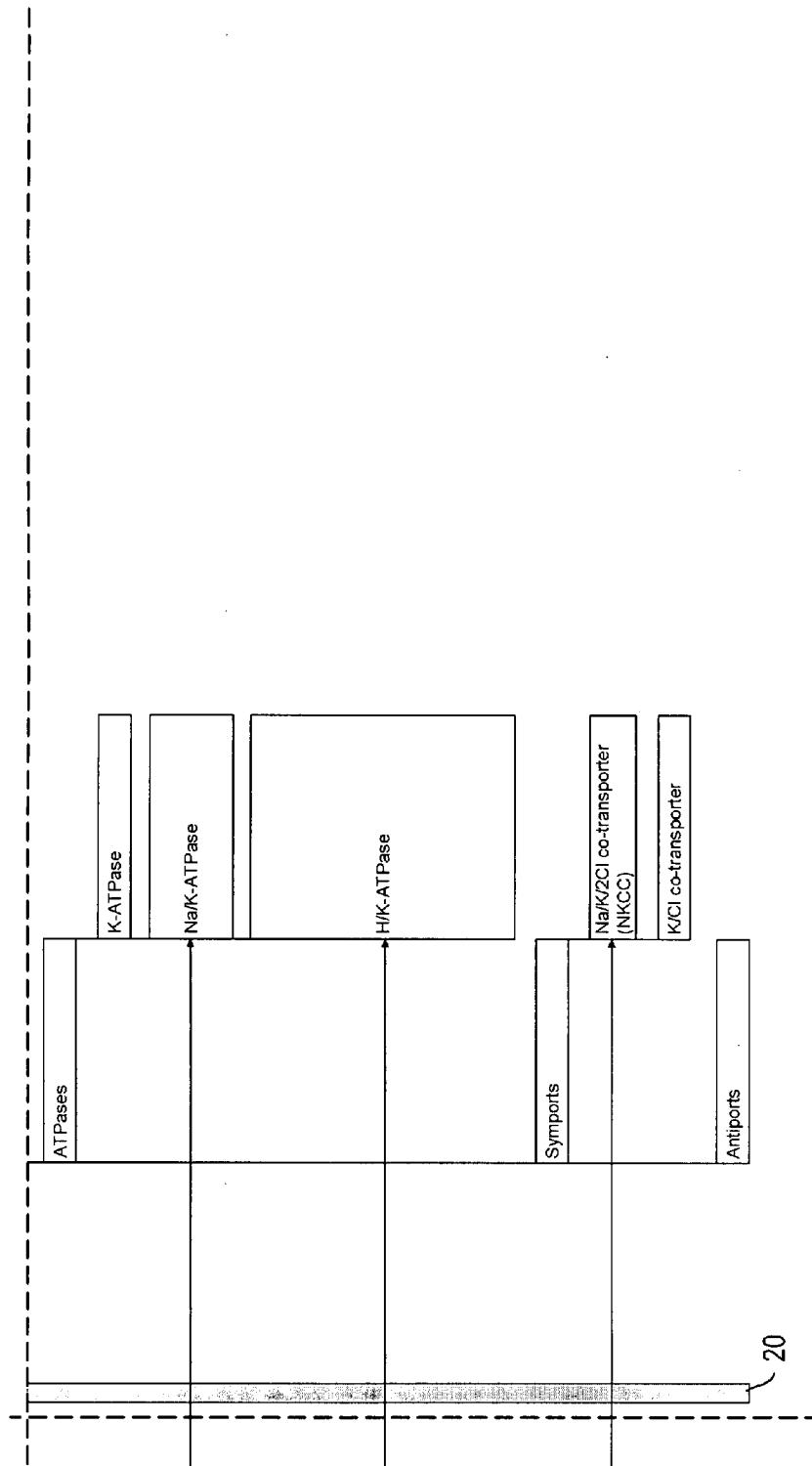


Figure 2J

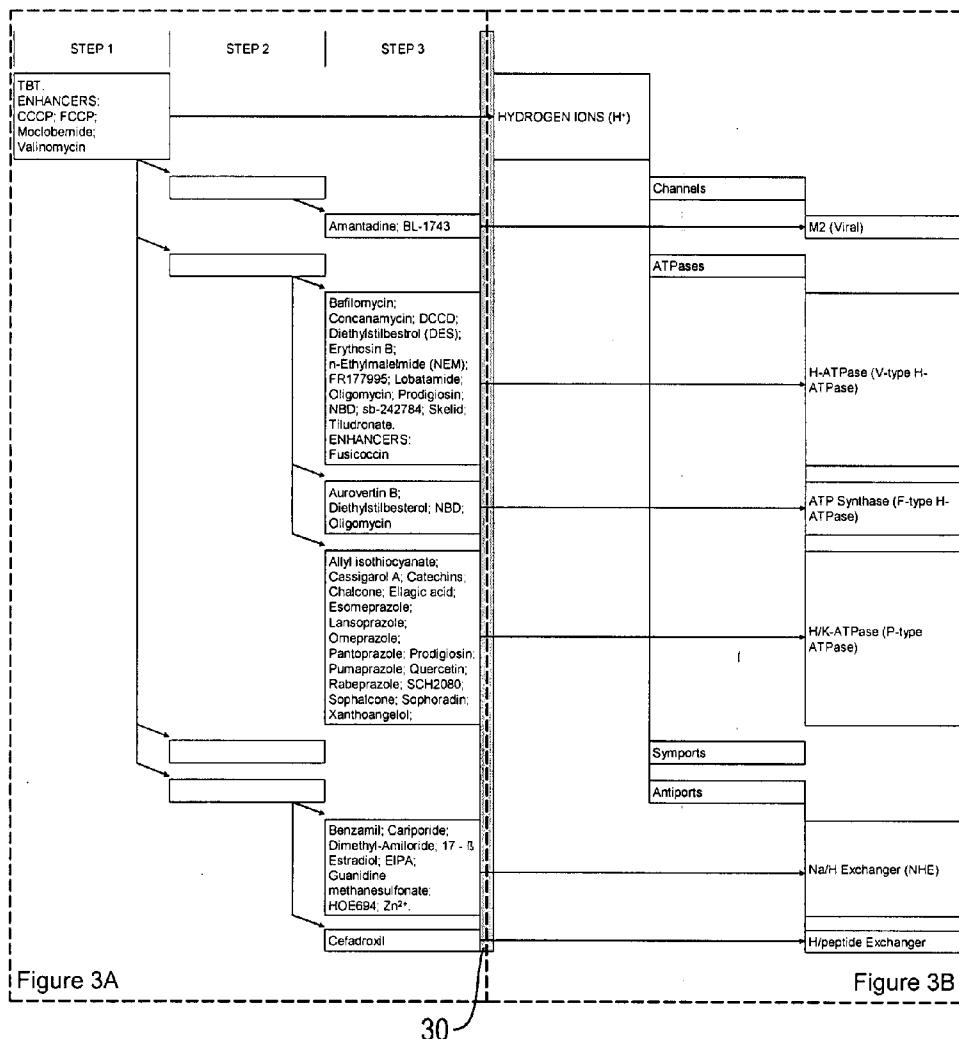


Figure 3

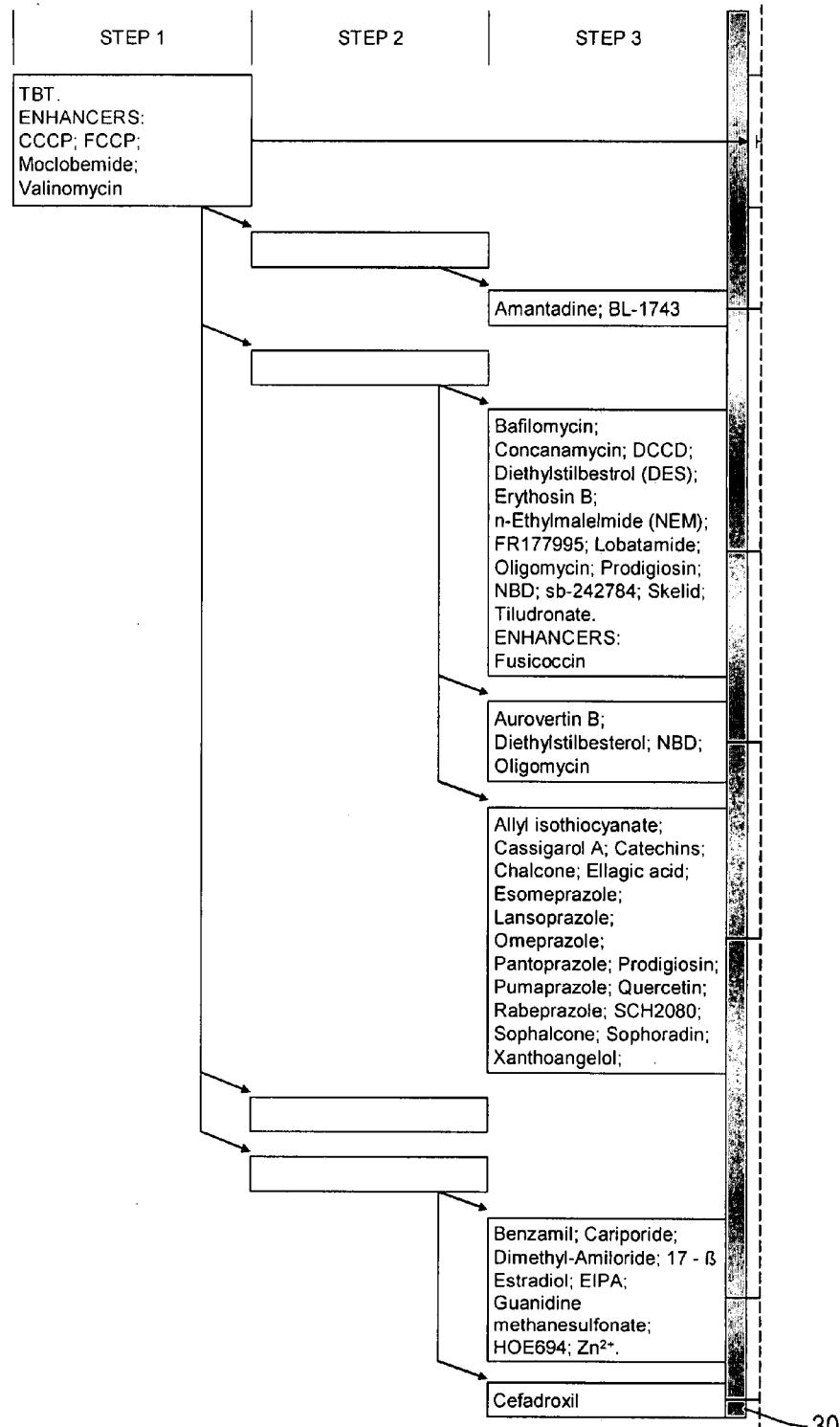
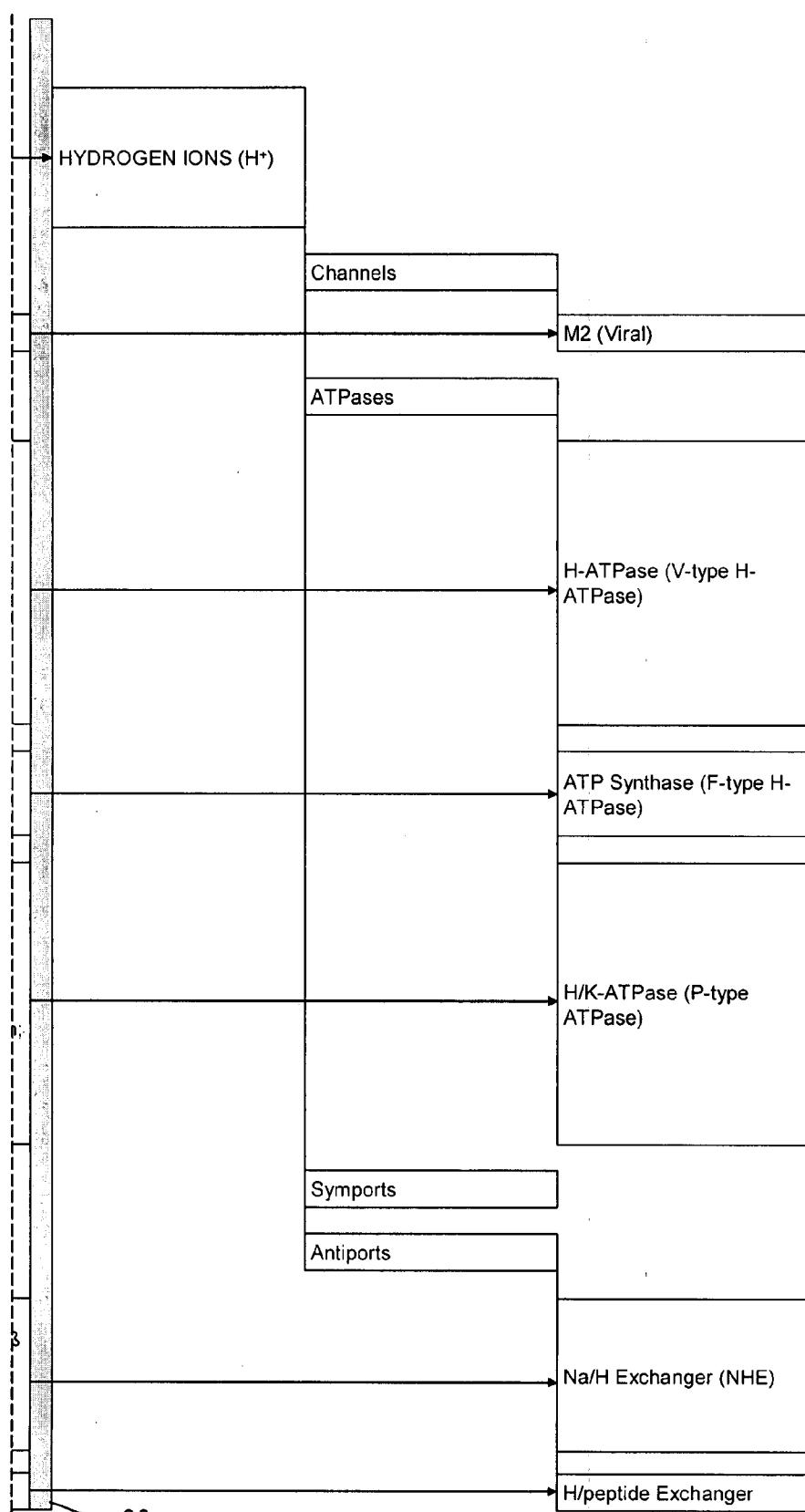
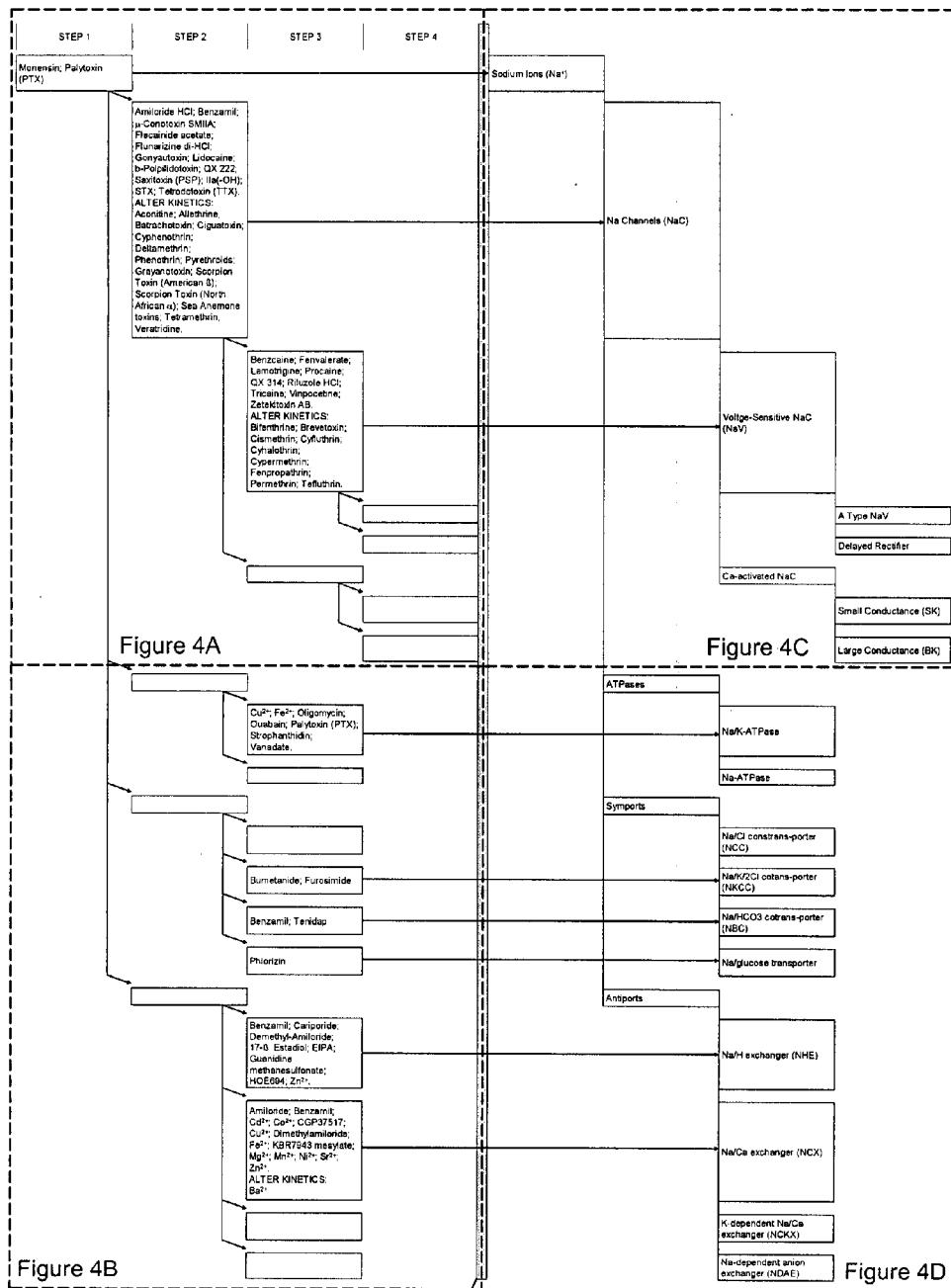


Figure 3A





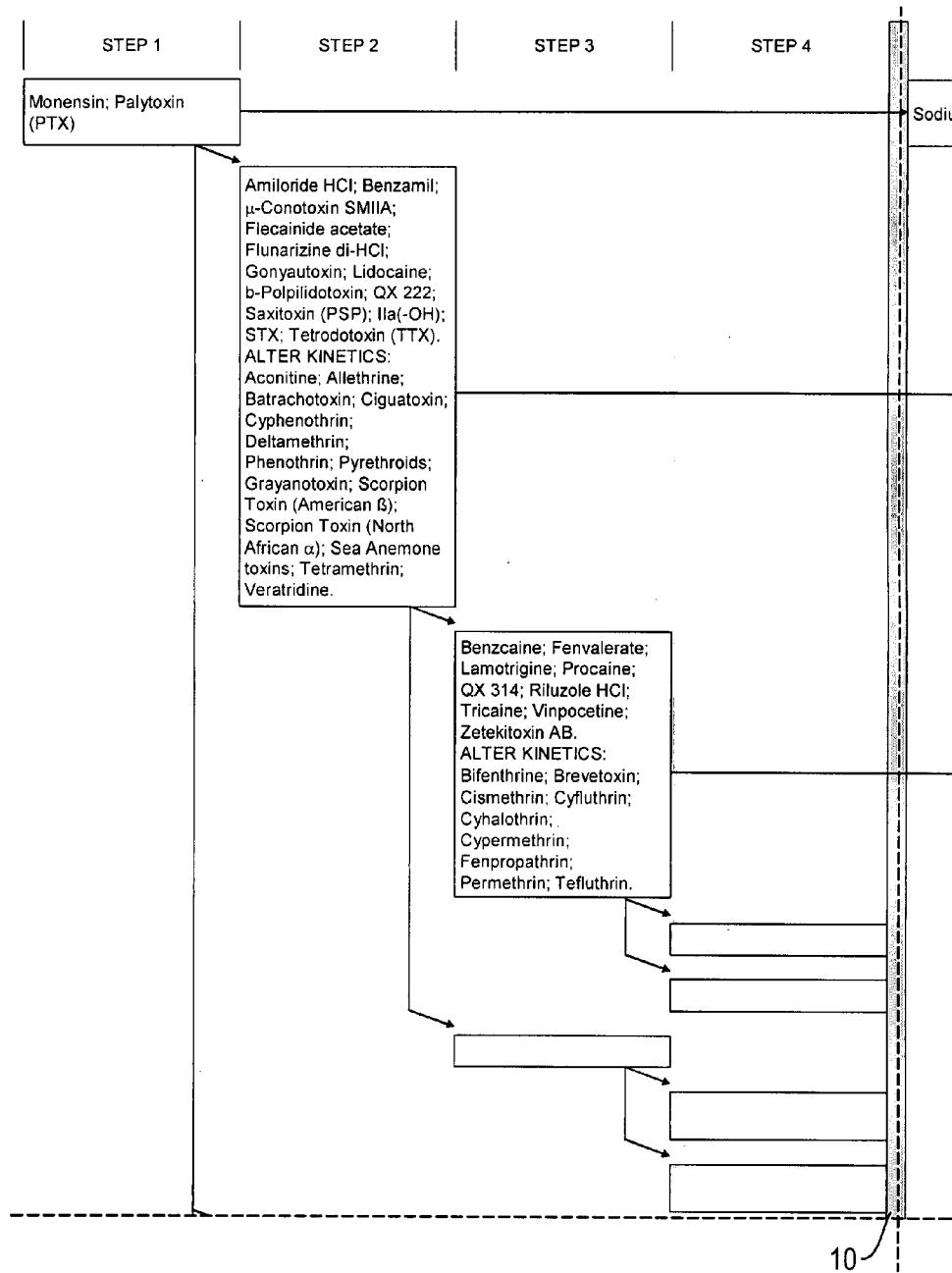


Figure 4A

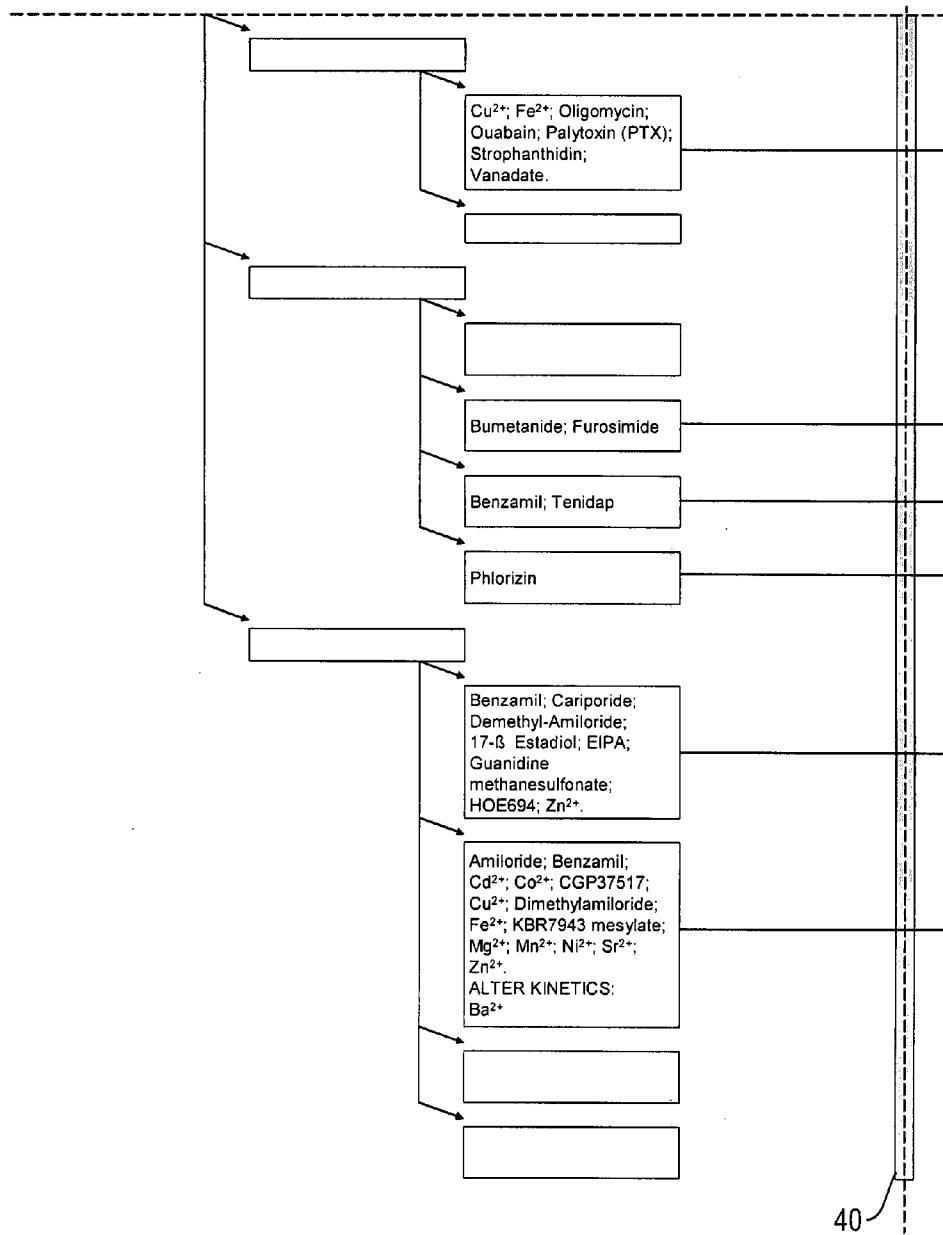


Figure 4B

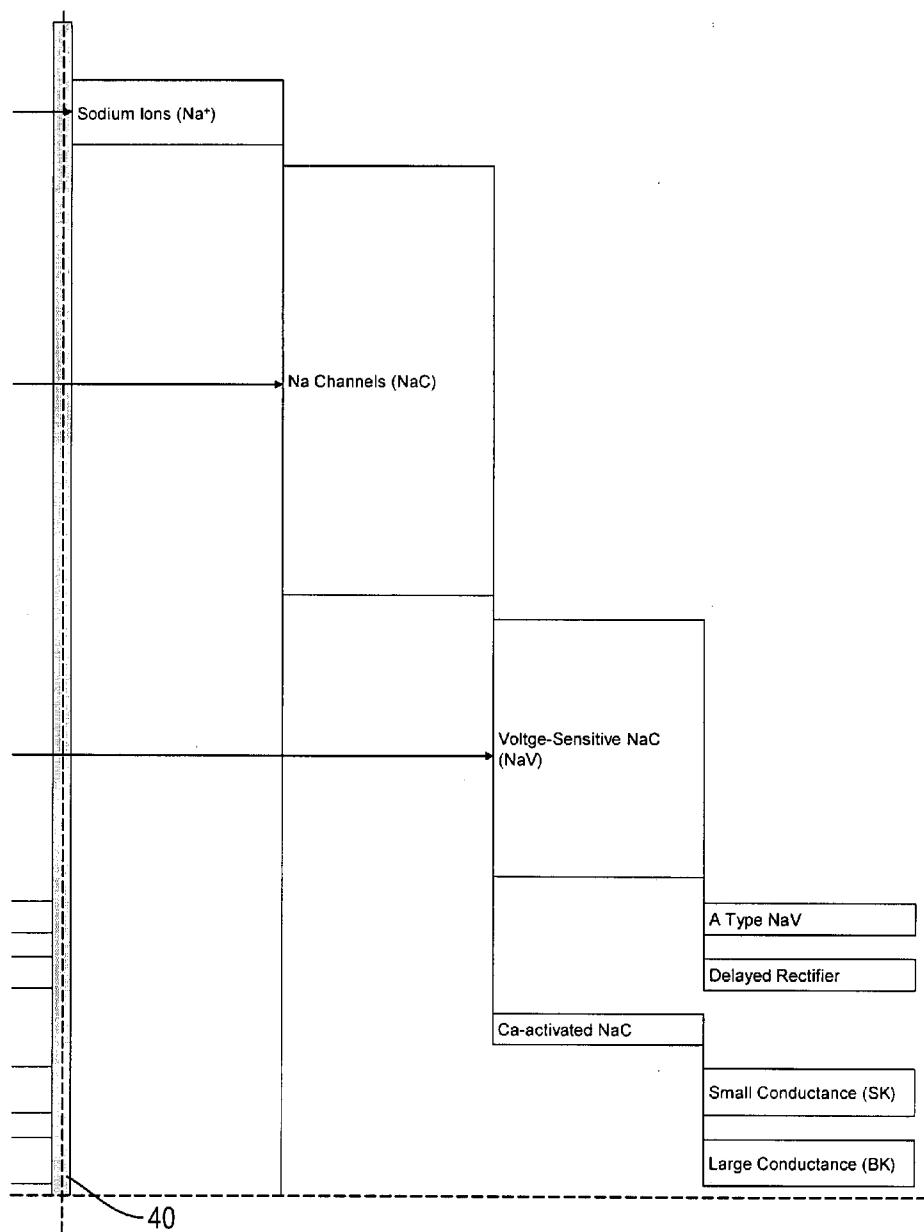


Figure 4C

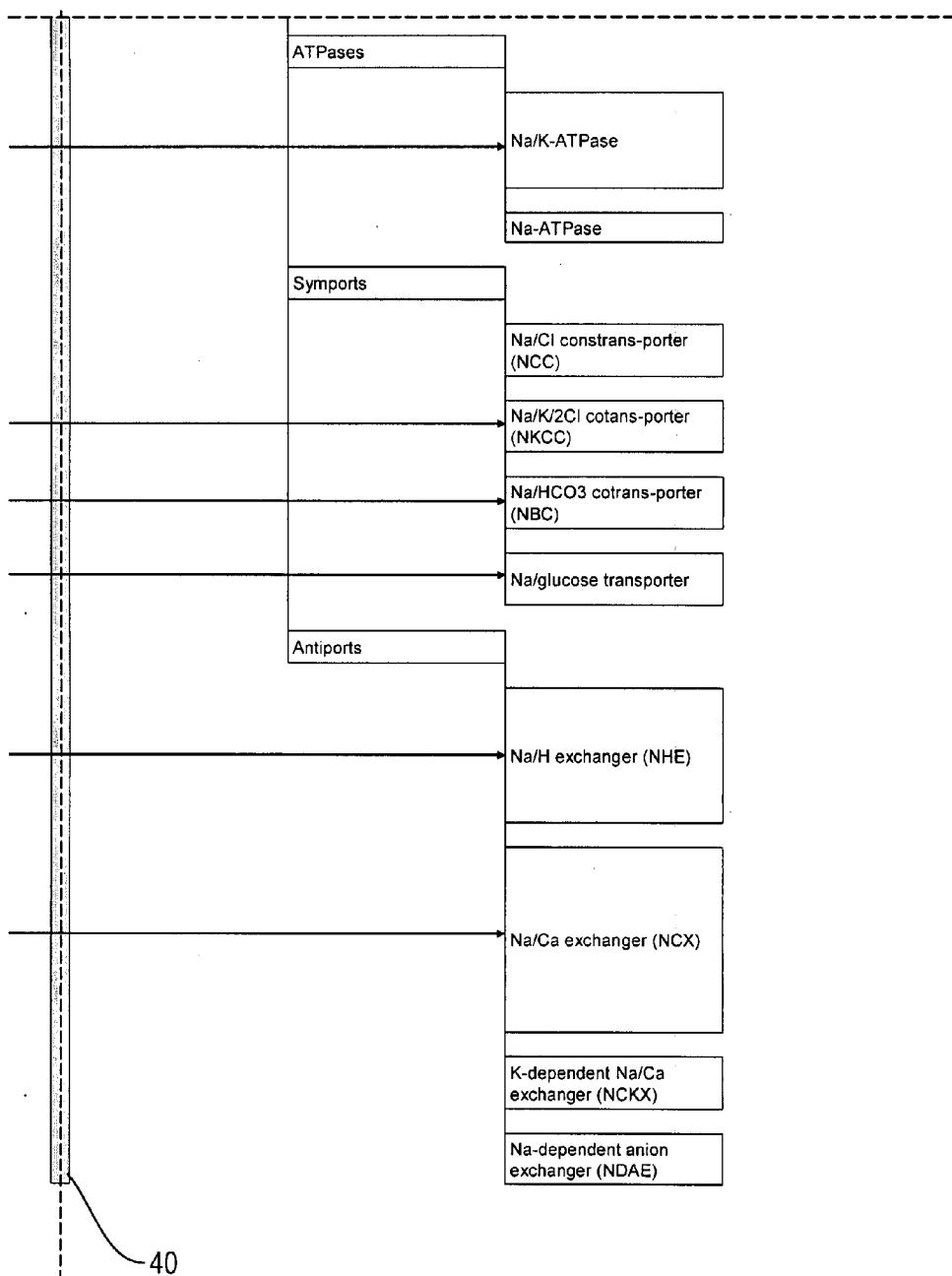


Figure 4D

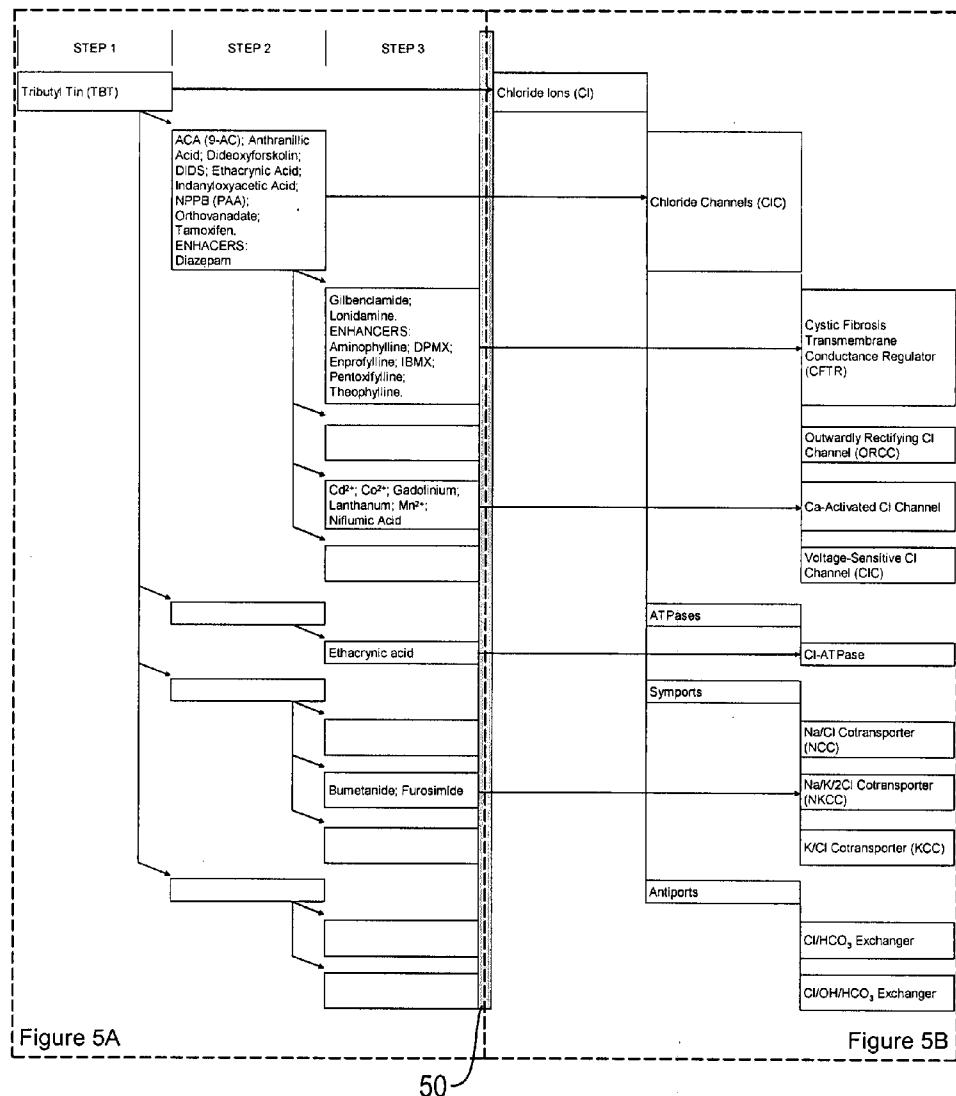


Figure 5

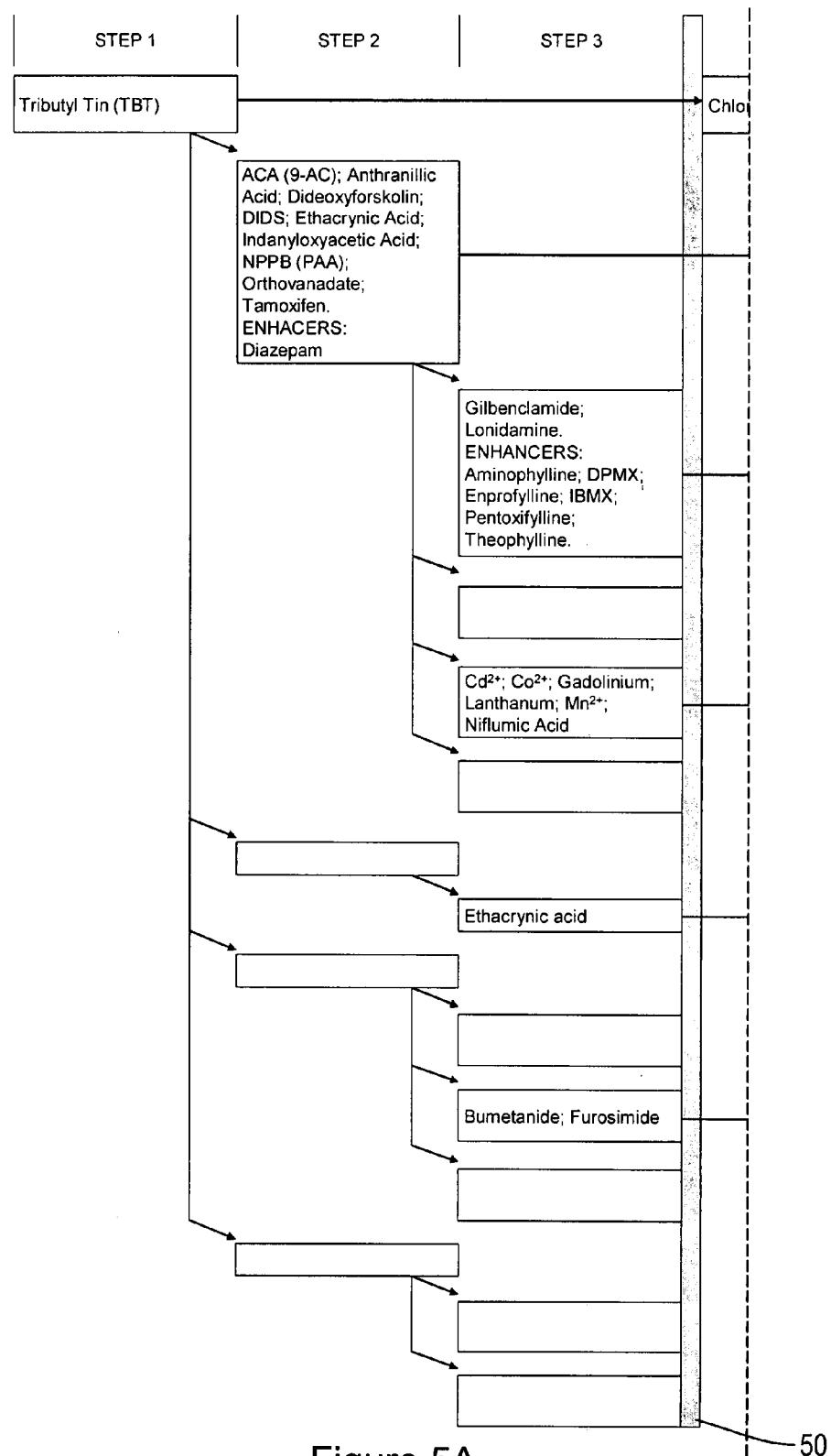


Figure 5A

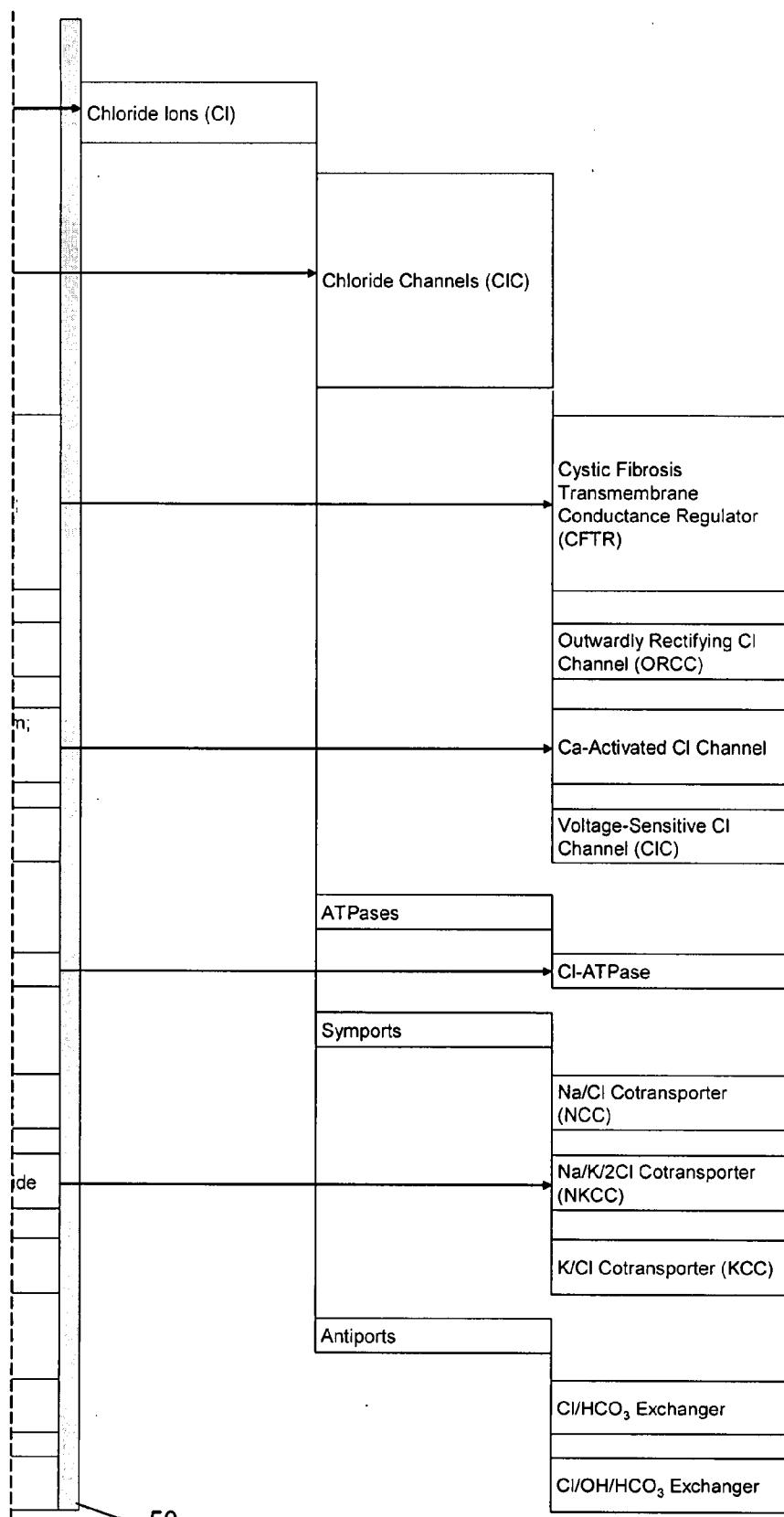


Figure 5B

DRUG	TARGETS
Cu ²⁺	Aquaporins
Trifluoroperazine	Calmodulin
Acetazolamide; Dichlorphenamide; Ethoxzolamide; Methazolamide.	Carbonic Anhydrase
Carbenoxolone; 18- β -Glycyrrhetic acid; Heptanol;Lindane; TAC-101	Gap Junctions (GJ)
Amitriptyline; Chloroquine;CuCl ² ;EPA; HOCl;Sodium Caprate	Tight Junctions
Streptolysin O.	Membrane Integrity
Diltiazem;Nifedipine; Tetracaine;LY83583	CNG (cyclic nucleotide gated) channels (monovalent cations, Ca ²⁺)
Cs ⁺ ;ZD7288; Ivabradine;Zatebradine; Alnidine	HCN (hyperpolarisation-activated cyclic nucleotide gated) channels (Na, K)
Gadolinium	Mechanosensitive channels
TPEN	Zn ²⁺ ,Cu ²⁺ ,Fe ²⁺ (ignores Ca ²⁺ and Mg ²⁺)

Figure 6

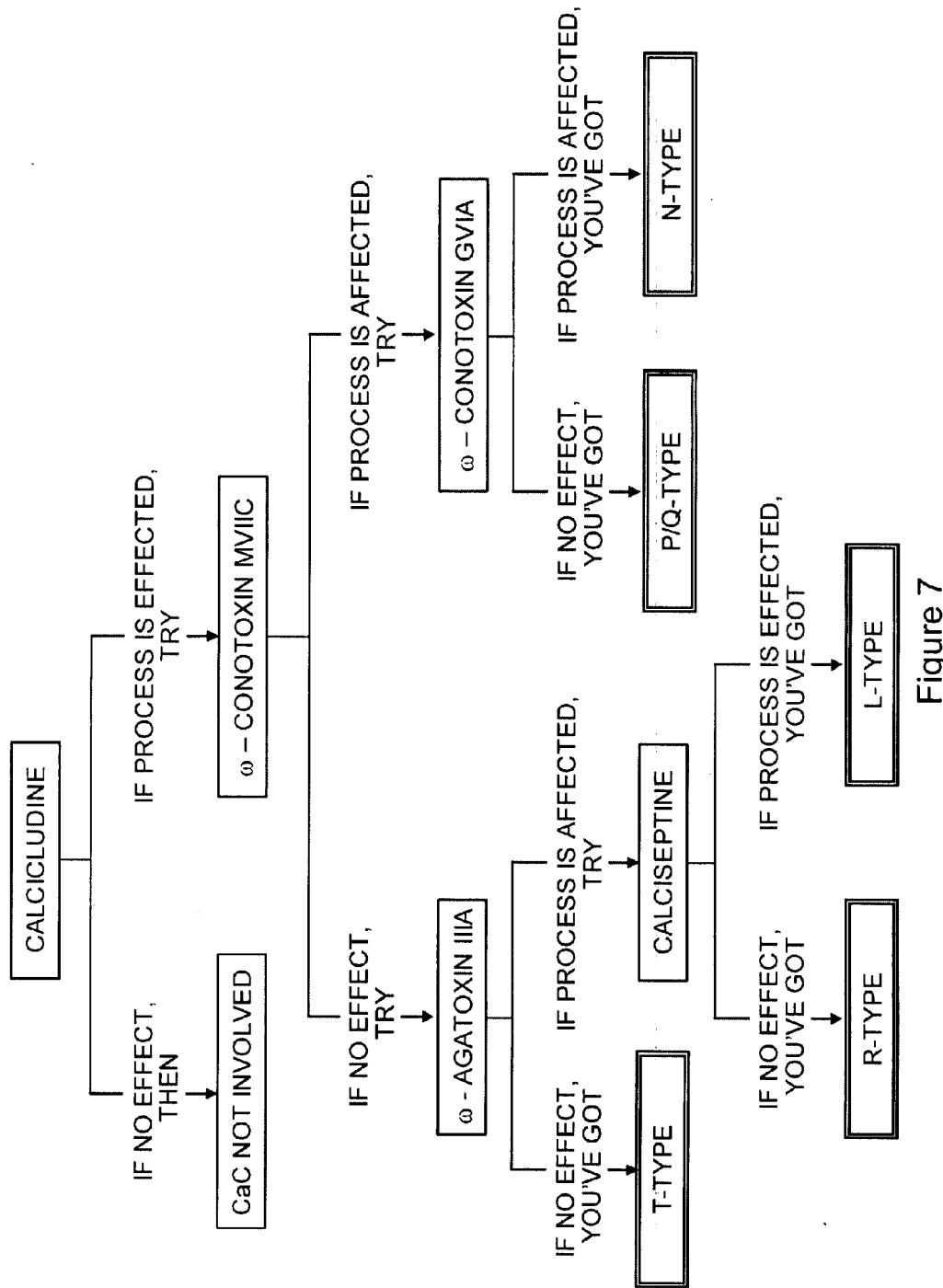


Figure 7

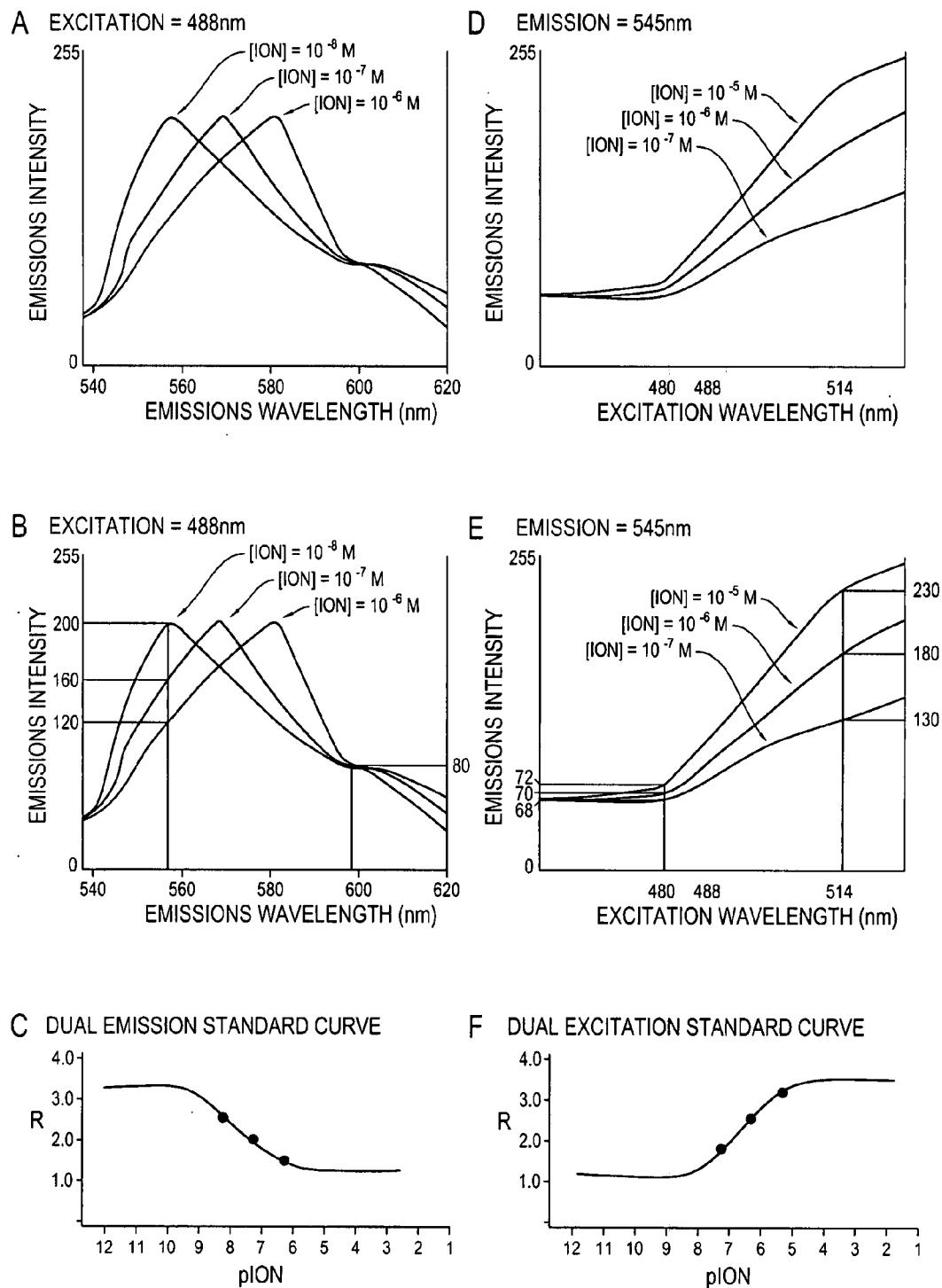


Figure 8

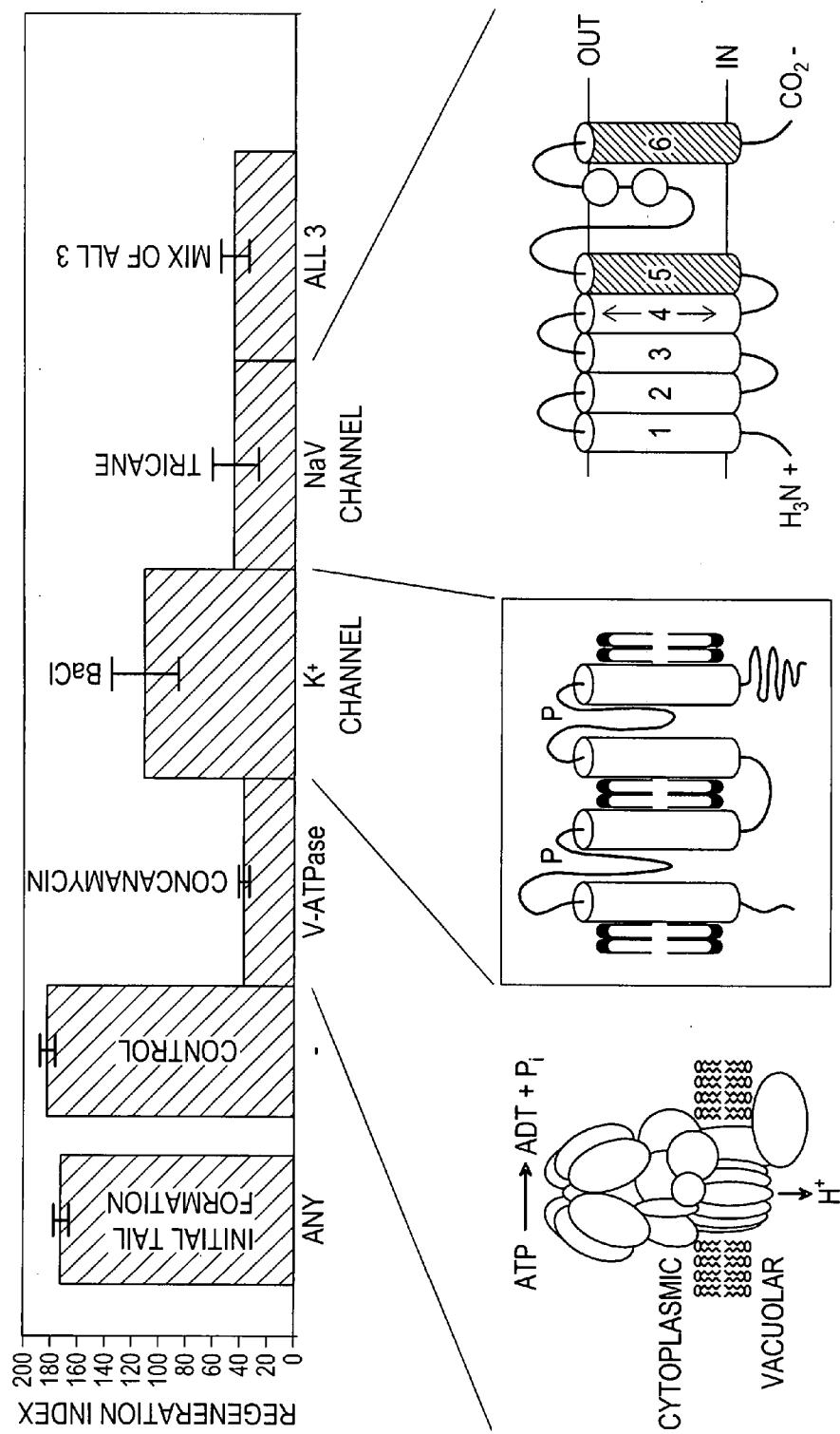


Figure 9

Figure 10A



Figure 10B

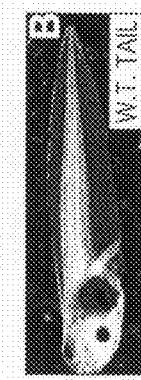


Figure 10C

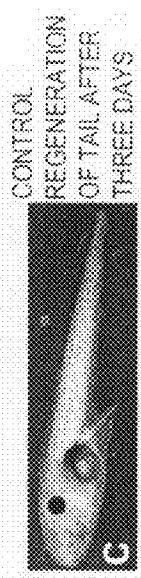


Figure 10D

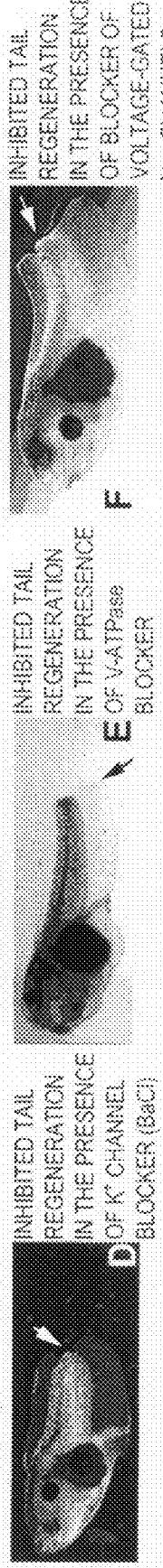


Figure 10E

No CHANNELS

Figure 10F

No CHANNELS

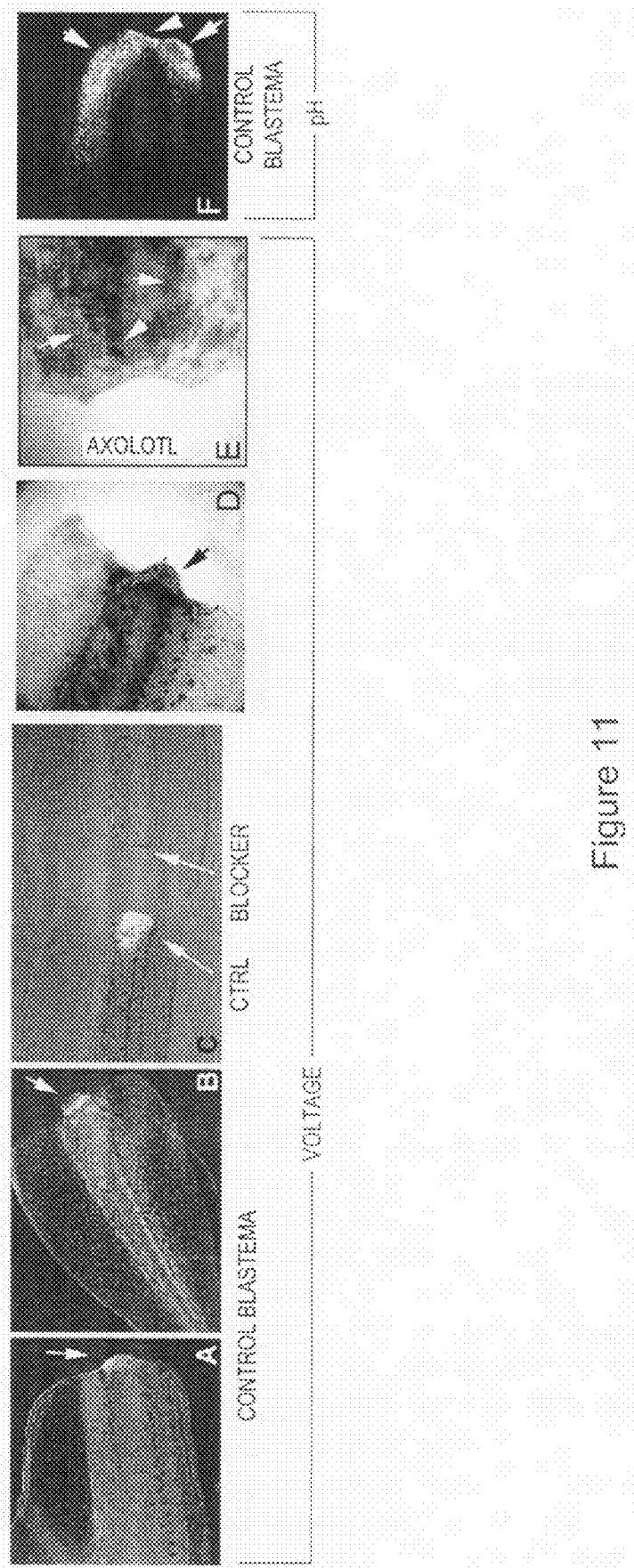
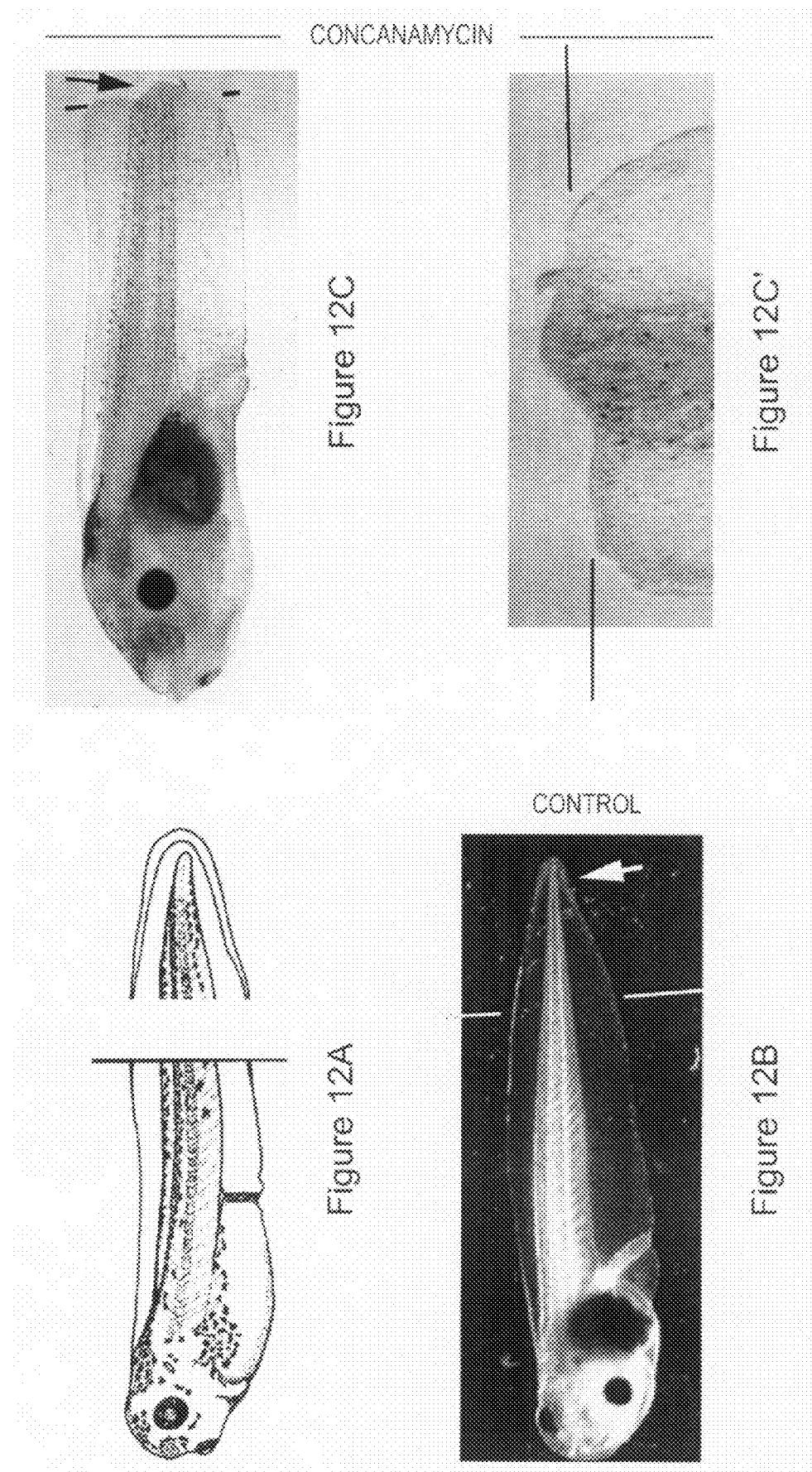
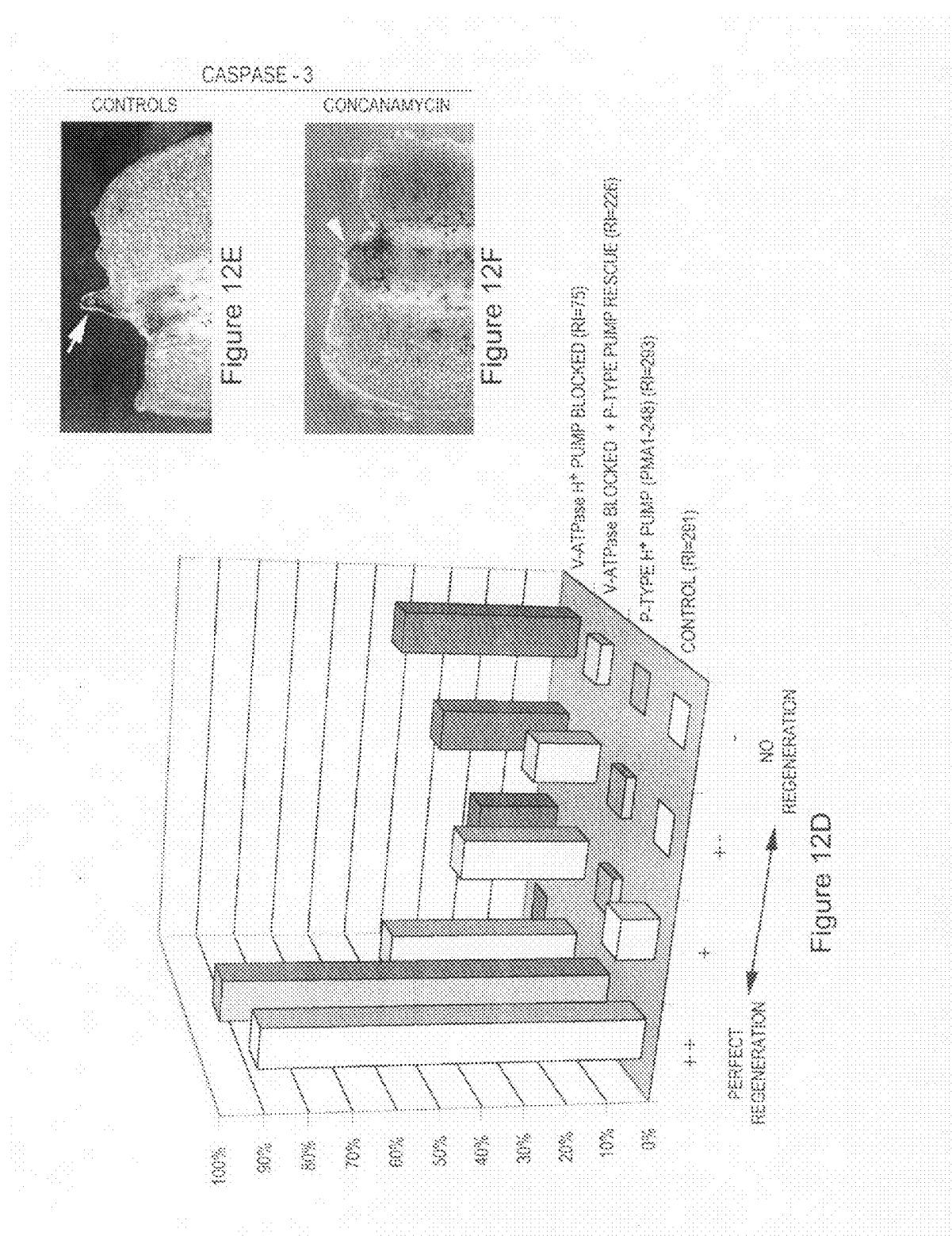


Figure 11





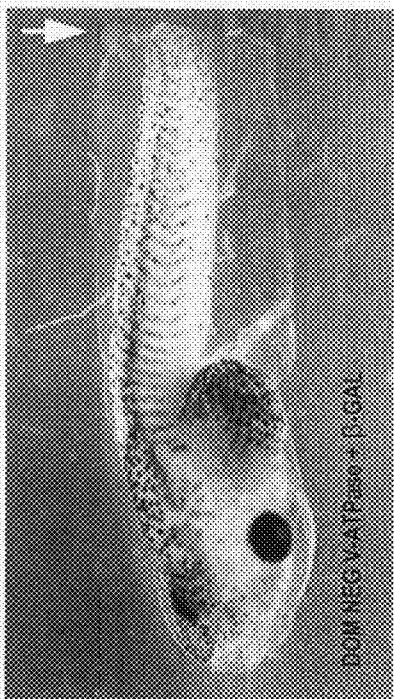


Figure 12G

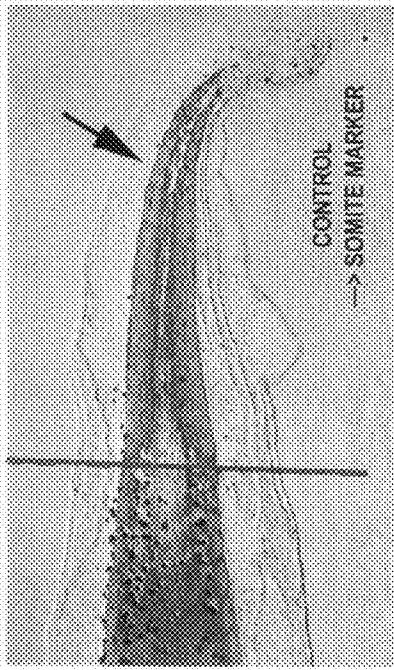


Figure 12H

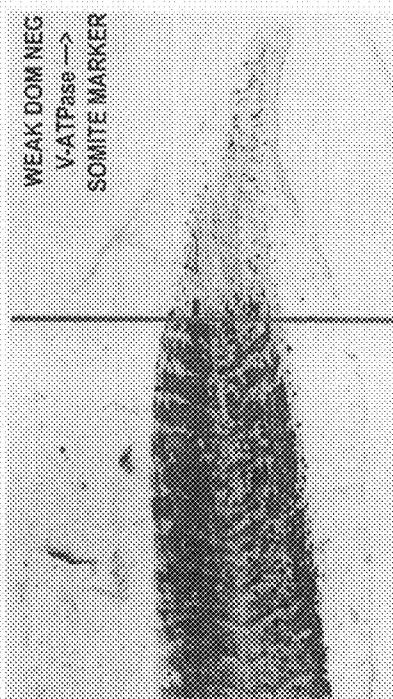


Figure 12I

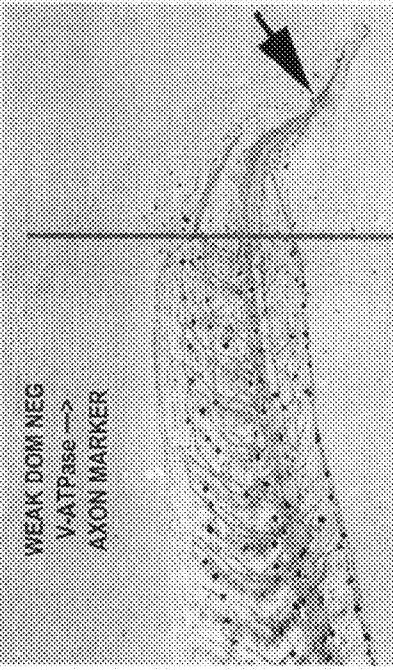


Figure 12J

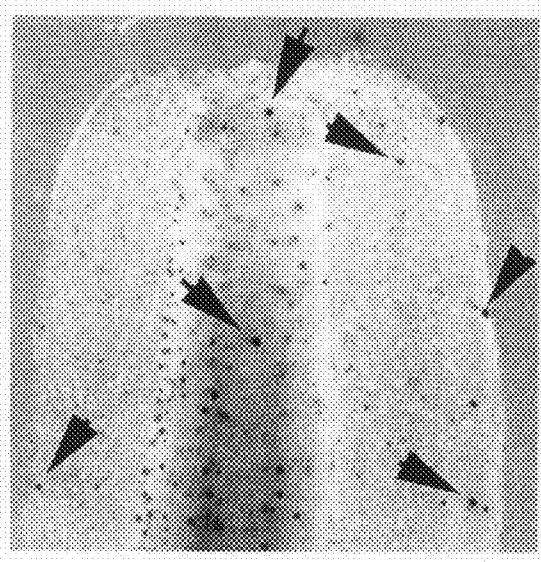


Figure 13A

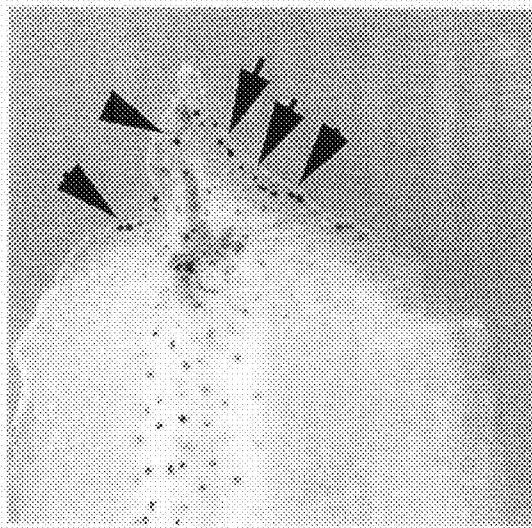


Figure 13B

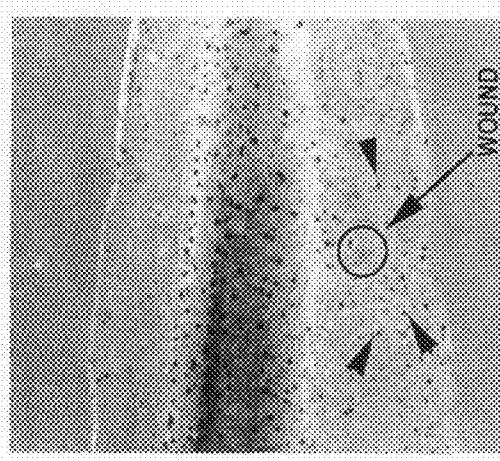


Figure 13C

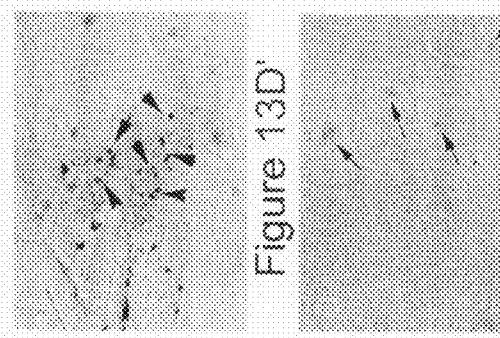
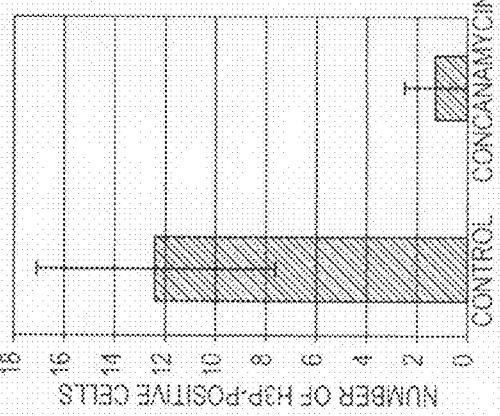
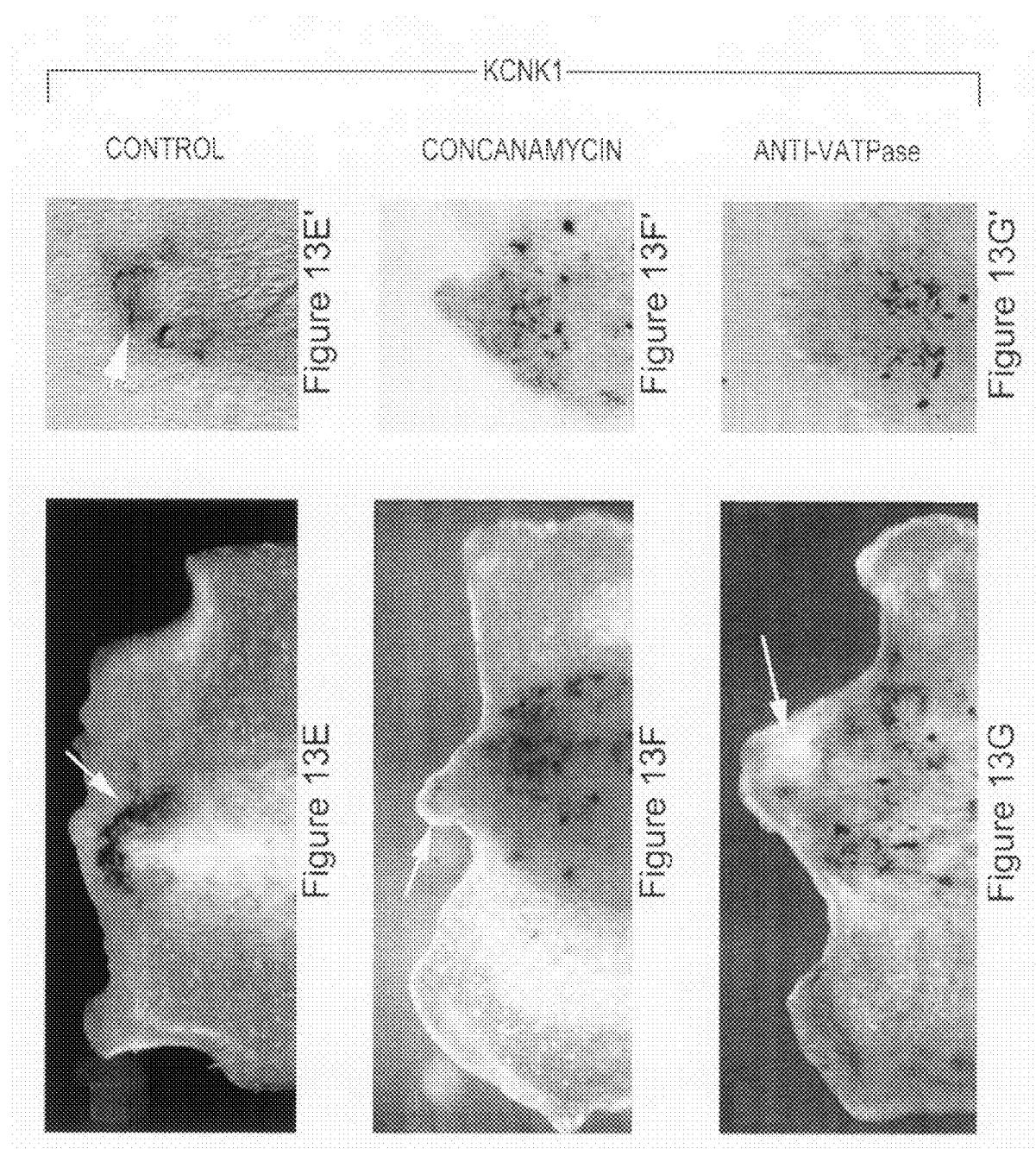
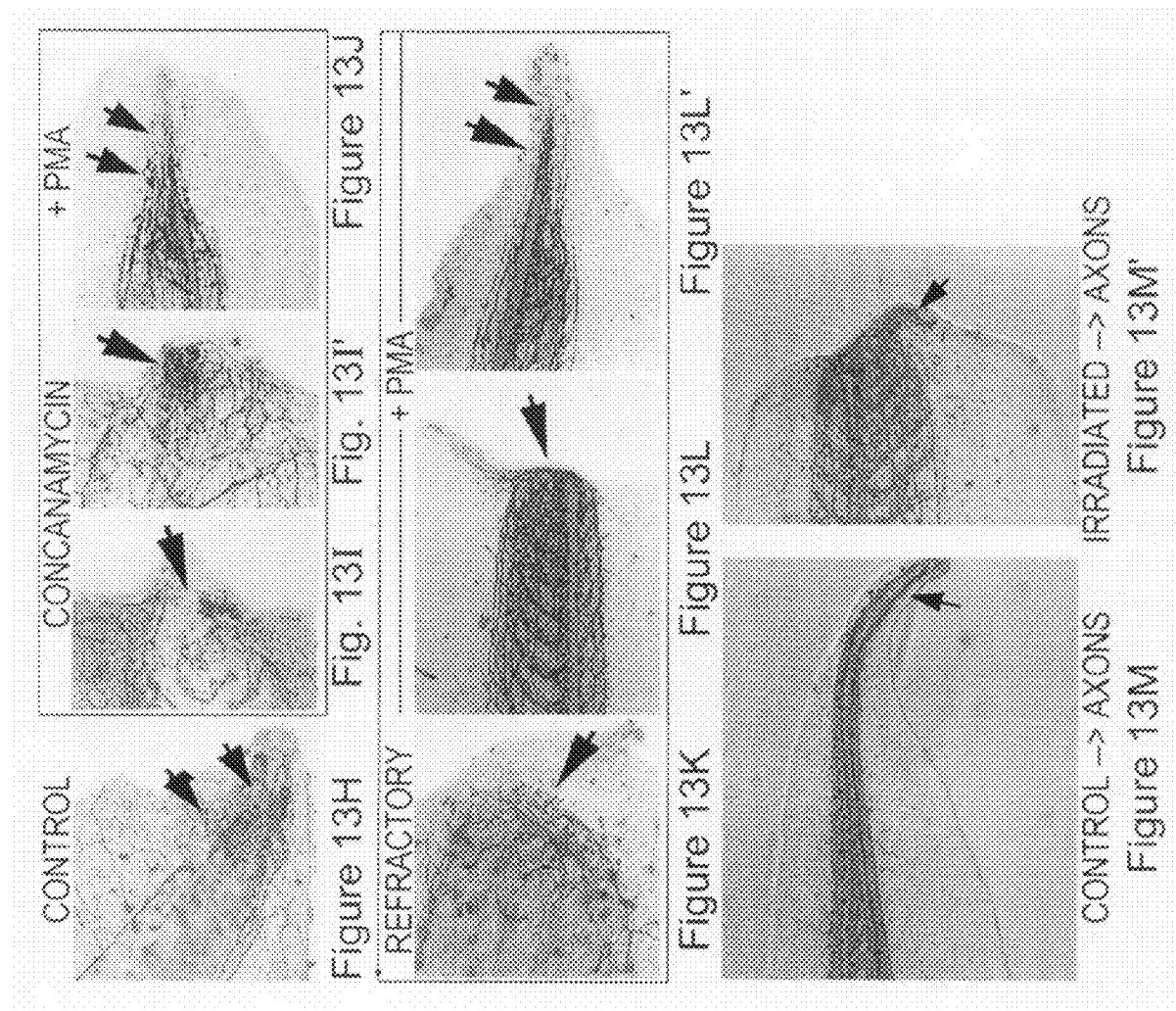


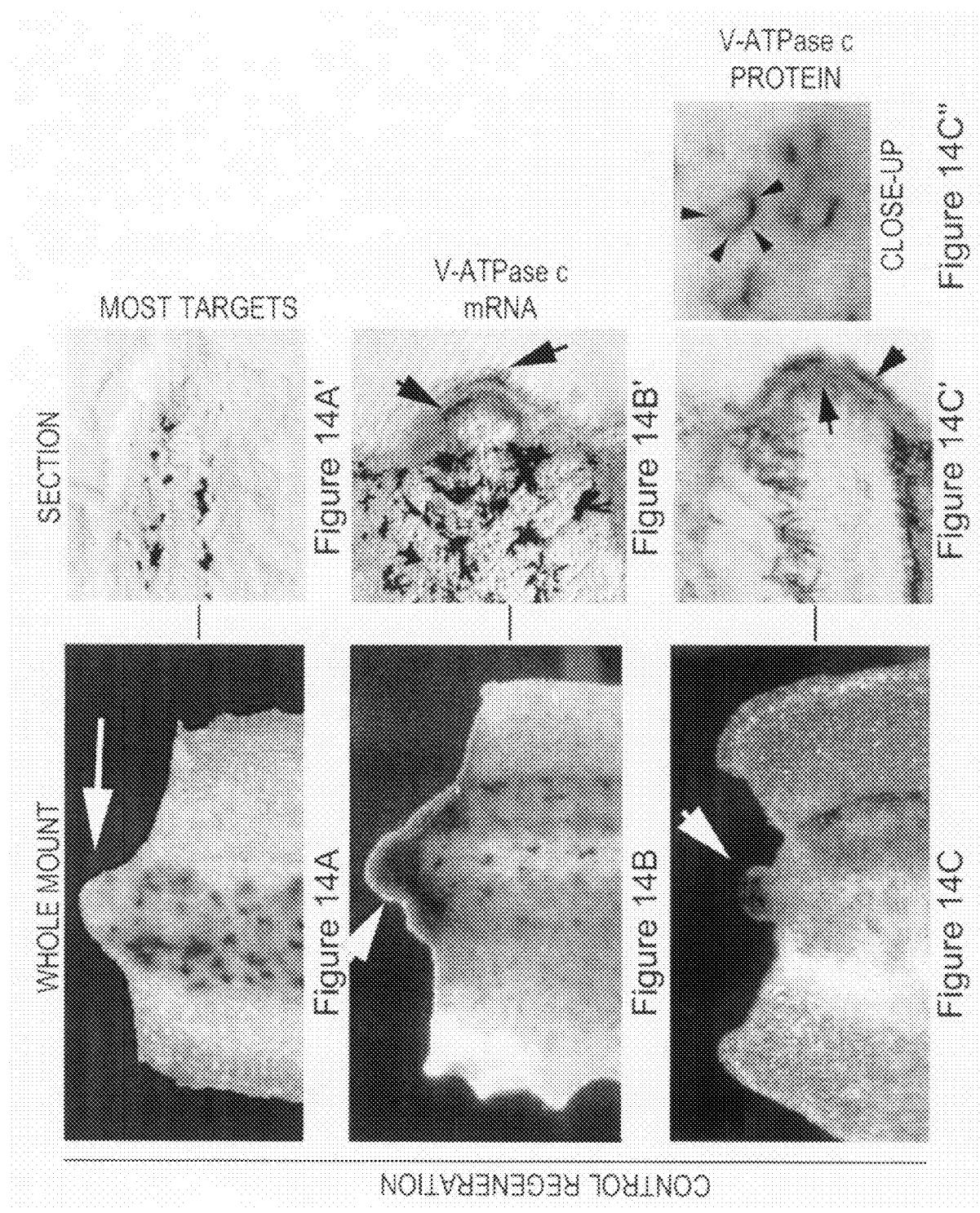
Figure 13D'



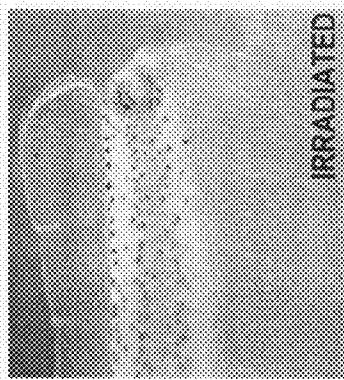
Figure 13D"







CELL PROLIFERATION (H3P)



REFRACTORY

IRRADIATED

Figure 14F

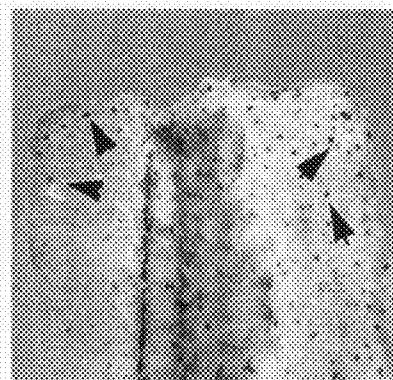


Figure 14G

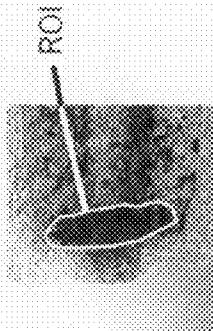


Figure 14H

V-ATPase c PROTEIN



REFRACTORY

IRRADIATED

Figure 14D



Figure 14E

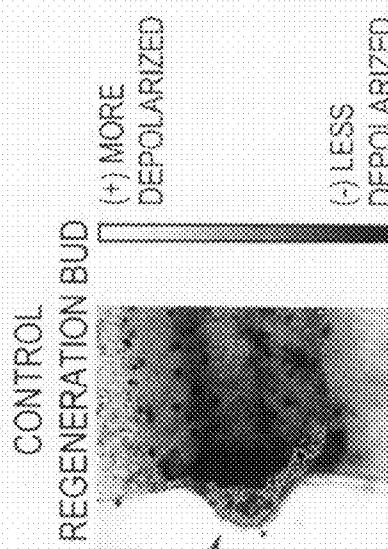


Figure 14H

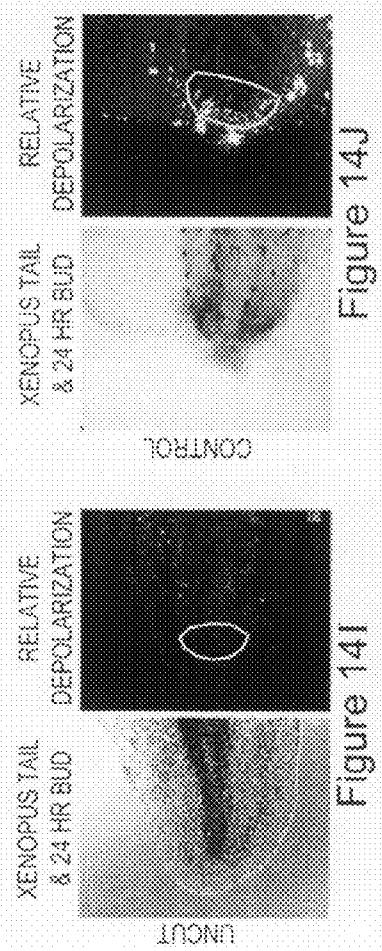


Figure 14J

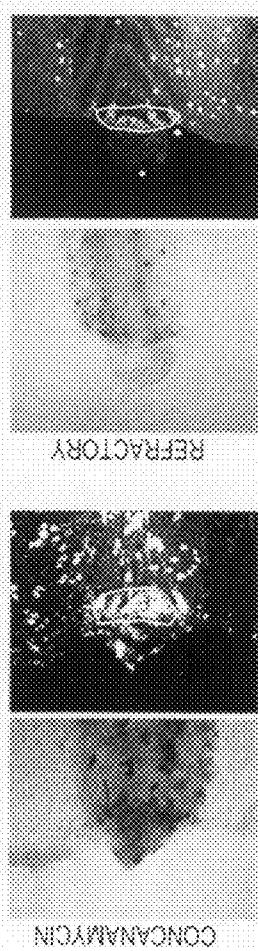


Figure 14L



Figure 14N

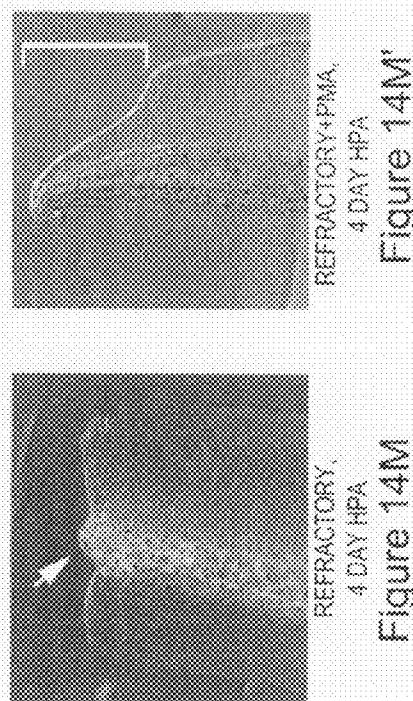


Figure 14M"

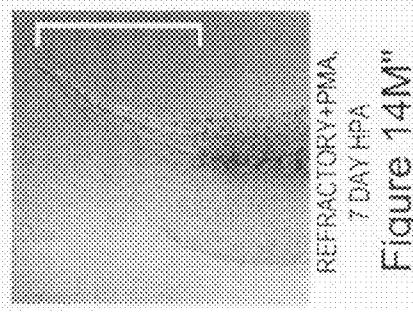


Figure 14N"

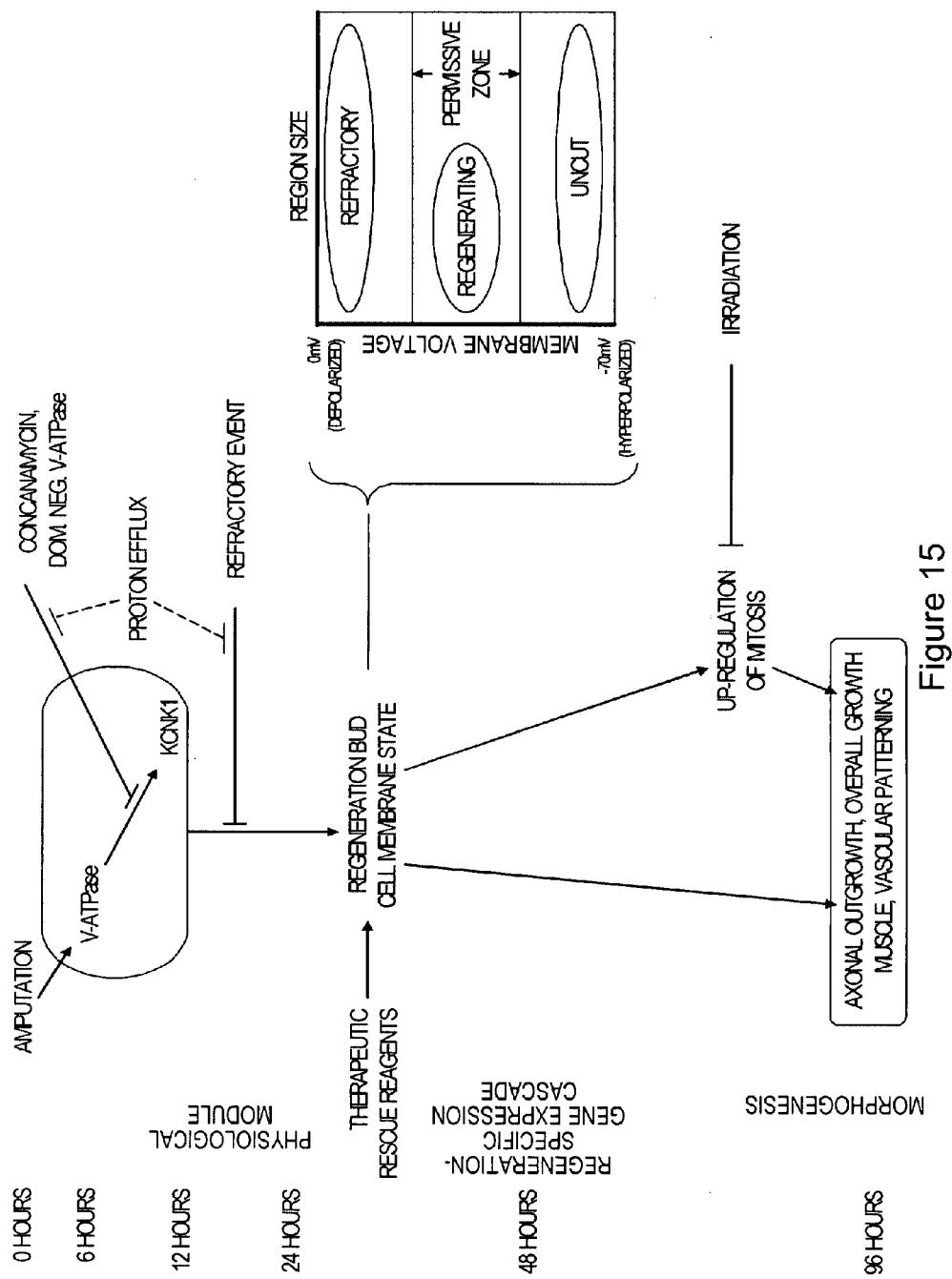


Figure 15

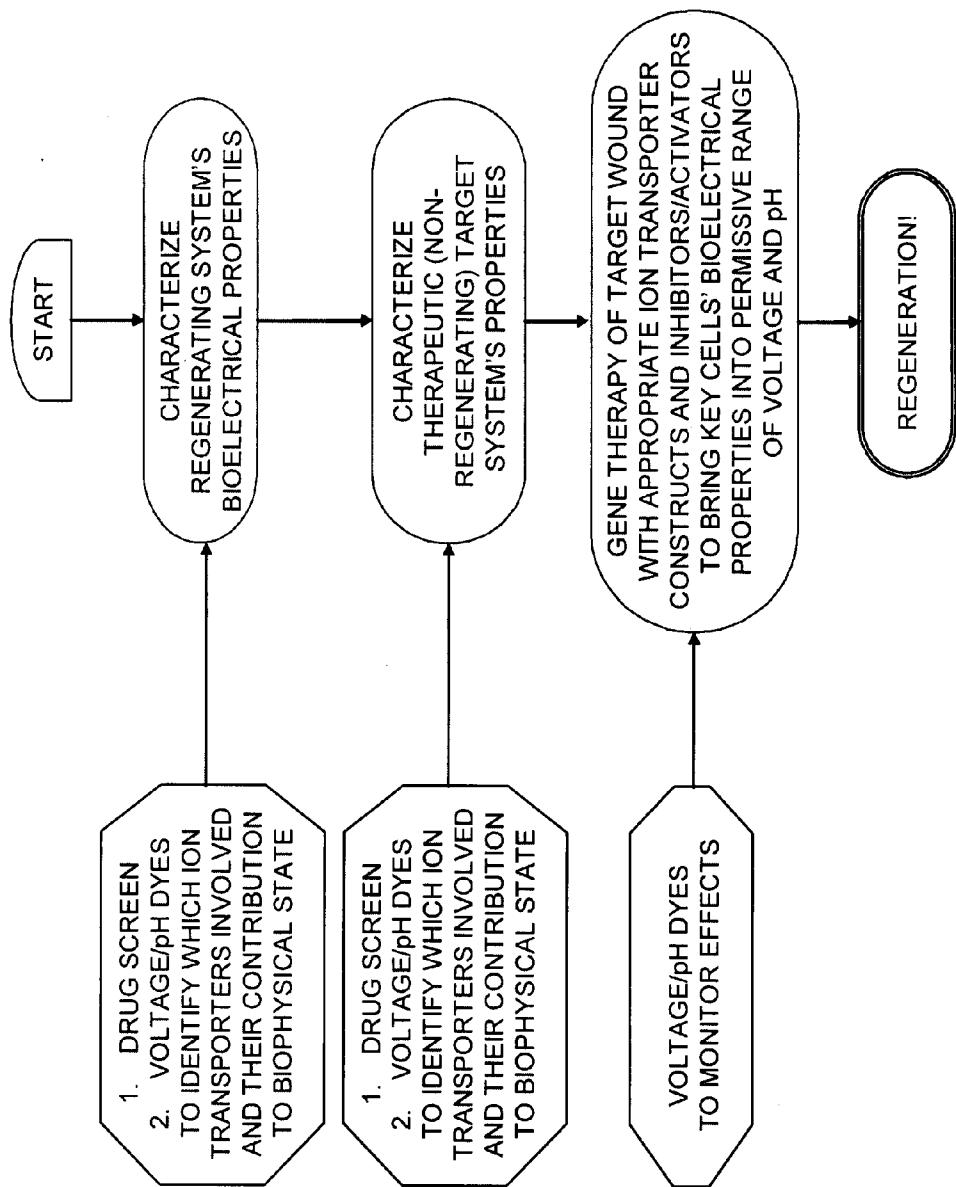


Figure 16

ION FLUX IN BIOLOGICAL PROCESSES, AND METHODS RELATED THERETO

RELATED APPLICATIONS

[0001] This application claims the benefit of priority to U.S. provisional application Ser. No. 60/841,777, filed Aug. 31, 2006, and U.S. provisional application Ser. No. 60/723,414, filed Oct. 4, 2005. The foregoing disclosures are hereby incorporated by reference in their entirety.

FUNDING

[0002] The invention described herein was supported, in whole or in part, by the National Center for Research Resources, National Institute of Health, Research Facilities Improvement Grant Number grant no. CO6RR11244, the National Science Foundation Career grant IBM #0347295, National Institutes of Health training grant 1-T32-DE-08327, National Institute of Health grant 1-R01-GM-06227, and the American Cancer Society Research Scholar Grant RSG-02-046-01-DDC. The United States government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0003] In the past several decades, scientific advances have enhanced our understanding of the molecular and cell biological basis for embryonic and adult development. However, the biophysics of development has received less attention. As such, a complete understanding of the role of biophysics, and the integration of biophysical, molecular, and cell biological events during development and disease remains lacking.

[0004] Cells are bounded by cell membranes. One important biophysical event that occurs throughout the development and life of cells results from the function of ion transporter proteins. Ion flows set up by ion transporter proteins such as ion channels, ion pumps, and gap junctions produce pH and voltage gradients within cells and across cell fields. The membrane potential and ion flux across cell membranes is crucial not only as a means of regulating cellular homeostasis, but also to help mediate specific biological processes.

[0005] Despite the importance of ion transporter proteins in development, the prior art fails to provide methods for effectively identifying what, if any, role ion transporters play in particular biological processes. As such, the prior art fails to provide guidance to move beyond a general appreciation that ion transporters may be important. The present invention provides methods for efficiently and effectively identifying particular roles for ion transporter proteins during embryonic and adult development or disease.

[0006] Once a role during a particular biological process is established, modulation of ion flux and/or expression of ion transporter proteins can be used to specifically regulate that particular biological process. The present invention provides examples whereby ion flux can be used to promote or inhibit a particular biological process.

BRIEF SUMMARY OF THE INVENTION

[0007] The present invention provides methods for promoting dedifferentiation and/or regeneration by modulating membrane potential and/or intracellular pH in cells.

[0008] In a first aspect, the invention provides a method of promoting regeneration, comprising providing a population of naturally regenerating cells; determining the membrane potential or pH range permissive for regeneration in said

population of regenerating cells; providing a population of non-regenerating cells; determining the membrane potential or pH range of said population of non-regenerating cells; and contacting the population of non-regenerating cells with an agent that modulates ion flux mediated by a class of ion transporter proteins. The agent modifies the membrane potential or pH of one or more cells in the population of non-regenerating cells to the range permissive for regeneration as determined in the second step of the method; thereby promoting regeneration of one or more cells in the population of cells.

[0009] In a second aspect, the invention provides a method of promoting regeneration, comprising determining the membrane potential or pH range permissive for regeneration in a population of regenerating cells; providing a population of non-regenerating cells; determining the membrane potential or pH range of said population of non-regenerating cells; and contacting the population of non-regenerating cells with an agent that modulates ion flux mediated by a class of ion transporter proteins. The agent modifies the membrane potential or pH of one or more cells in the population of non-regenerating cells to the range permissive for regeneration as determined in the second step of the method; thereby promoting regeneration of one or more cells in the population of cells.

[0010] In one embodiment of any of the foregoing, determining the membrane potential or pH range permissive for regeneration in said population of regenerating cells comprises providing a population of regenerating cells; contacting said population of cells with an agent which is a voltage sensitive agent that produces a detectable signal; and measuring the detectable signal to calculate an average membrane potential or pH of said population of cells during regeneration. In another embodiment of any of the foregoing, the first two steps are repeated for multiple types of naturally regenerating cells and the average membrane potential or pH range are used in the final step. In another embodiment of any of the foregoing, determining the membrane potential or pH range of said population of non-regenerating cells comprises providing a population of non-regenerating cells; contacting said population of cells with an agent which is a voltage sensitive agent that produces a detectable signal; and measuring the detectable signal to calculate an average membrane potential or pH of said population of cells.

[0011] In one embodiment of any of the foregoing, the method is an in vitro method and the population of non-regenerating cells are in culture. In another embodiment of any of the foregoing, the method is an in vitro method and the population of non-regenerating cells are in a preparation of tissue in culture. In another embodiment of any of the foregoing, the method is an in vivo method and the population of non-regenerating cells are in an animal.

[0012] In one embodiment of any of the foregoing, the population of non-regenerating cells are derived from or resident in an animal, and the animal is a flatworm, an amphibian, a fish, a reptile, a bird, or a mammal. In another embodiment of any of the foregoing, the mammal is a mouse, rat, cat, dog, rabbit, goat, hamster, pig, sheep, non-human primate, or human. In another embodiment of any of the foregoing, the mammal is a human. In another embodiment of any of the foregoing, the human is a patient in need of regeneration. In another embodiment of any of the foregoing, the population of non-regenerating cells are embryonic, fetal, larval, juve-

nile, or adult cells. In another embodiment of any of the foregoing, the cells are adult cells.

[0013] In one embodiment of any of the foregoing, the agent inhibits ion flux mediated by the class of transporter proteins. In another embodiment of any of the foregoing, the agent promotes ion flux mediated by the class of transporter proteins. In another embodiment of any of the foregoing, the agent is selected from a nucleic acid, a peptide, a protein, a small organic molecule, a small inorganic molecule, an antisense oligonucleotide, an RNAi construct, or an antibody.

[0014] In one embodiment of any of the foregoing, the agent that inhibits ion flux mediated by a class of ion transporters is an ion channel protein or a nucleotide construct that encodes an ion channel protein. In another embodiment of any of the foregoing, the ion transporter protein is a hyperpolarizing transporter. In another embodiment of any of the foregoing, the ion transporter protein is a depolarizing transporter. In another embodiment of any of the foregoing, the ion transporter protein is an H⁺ pump. In another embodiment of any of the foregoing, the H⁺ pump is a V-ATPase H⁺ pump. In another embodiment of any of the foregoing, the H⁺ pump is a P-type H⁺ ATPase pump. In another embodiment of any of the foregoing, the H⁺ pump is a yeast PMA1.2H⁺ pump. In another embodiment of any of the foregoing, the ion transporter protein is a K⁺ channel. In another embodiment of any of the foregoing, the K⁺ channel is a ROMK K⁺ channel. In another embodiment of any of the foregoing, the K⁺ channel is an ERG K⁺ channel. In another embodiment of any of the foregoing, the ion transporter protein is an Na⁺ channel. Note, however, that these particular ion transporter proteins are merely exemplary of particular proteins that can be manipulated to modulate membrane potential. One of skill in the art can readily select an ion transporter protein and manipulate the activity of that transporter protein (using an agent) to depolarize or hyperpolarize cell membranes, thereby shifting membrane potential of cells or cells in a tissue into a range permissive for regeneration. In certain embodiments, the particular ion transporter protein is chosen because it is endogenously expressed in the particular cells or tissues being studied. In certain embodiments, the particular ion transporter protein is chosen because it is not endogenously expressed in the particular cells or tissues being studied.

[0015] In a third aspect, the invention provides a method of promoting regeneration, comprising providing a population of cells of known membrane potential or pH; and contacting the population of cells with an agent that modulates ion flux mediated by a class of ion transporter proteins. The agent modifies the membrane potential or pH of one or more cells in the population of cells to a range permissive for regeneration; thereby promoting regeneration of one or more cells in the population of cells.

[0016] In one embodiment of any of the foregoing, the membrane potential range permissive for regeneration is between -70 mV and 30 mV. In another embodiment, the membrane potential range permissive for regeneration is between -40 mV and 20 mV. In still another embodiment, the lower range of the membrane potential is -70 mV, -60 mV, -50 mV, -40 mV, -35 mV, -30 mV, -25 mV, -20 mV, -15 mV, -10 mV, -5 mV, or 0 mV and the upper range of the membrane potential is -30 mV, -20 mV, -10 mV, -5 mV, 0 mV, 5 mV, 10 mV, 15 mV, 20 mV, 25 mV, or 30 mV.

[0017] In one embodiment of any of the foregoing, the pH range permissive for regeneration is less than or equal to about 7.0, 6.9, 6.8, 6.7, 6.6, 6.5, 6.4, 6.3, 6.2, 6.1, or 6.0. In

another embodiment of any of the foregoing, both the membrane potential and the pH range permissive for regeneration may be used in any of the steps of the methods of the invention.

[0018] In a fourth aspect, the invention provides a method of promoting regeneration, comprising providing a population of cells; and contacting the population of cells with an agent that increases ion flux mediated by a V-ATPase H⁺ pump. The agent promotes relative hyperpolarization of cell membranes of one or more cells in the population of cells, thereby promoting regeneration of one or more cells in the population of cells.

[0019] In one embodiment of any of the foregoing, the agent is a nucleotide construct encoding the V-ATPase H⁺ pump.

[0020] In a fifth aspect, the invention provides an agent that modulates activity of a H⁺ pump in the manufacture of a pharmaceutical composition for promoting regeneration.

[0021] In one embodiment of any of the foregoing, the H⁺ pump is a P-type H⁺ ATPase pump. In another embodiment of any of the foregoing, the H⁺ pump is a V-ATPase H⁺ pump.

[0022] In a sixth aspect, the invention provides an agent that modulates the activity of a K⁺ channel in the manufacture of a pharmaceutical composition for promoting regeneration.

[0023] In one embodiment of any of the foregoing, the channel is a ROMK K⁺ channel. In another embodiment of any of the foregoing, the channel is an ERG K⁺ channel. In another embodiment of any of the foregoing, the channel is an KATP channel. In another embodiment of any of the foregoing, the channel is an KCNK1 channel. In another embodiment of any of the foregoing, the channel is an KCNQ1 channel.

[0024] In a seventh aspect, the invention provides an agent that activates an voltage-dependent sodium channel (Na^V) in the manufacture of a pharmaceutical composition for promoting regeneration.

[0025] Note, that although certain embodiments and aspects of the invention involve manipulating particular ion transporter proteins, these particular ion transporter proteins are merely exemplary of particular proteins that can be manipulated to modulate membrane potential. One of skill in the art can readily select an ion transporter protein and manipulate the activity of that transporter protein (using an agent) to depolarize or hyperpolarize cell membranes, thereby shifting membrane potential of cells or cells in a tissue into a range permissive for regeneration. In certain embodiments, the particular ion transporter protein is chosen because it is endogenously expressed in the particular cells or tissues being studied. In certain embodiments, the particular ion transporter protein is chosen because it is not endogenously expressed in the particular cells or tissues being studied.

[0026] In an eighth aspect, the invention provides a method of screening for compounds that promote dedifferentiation of cells. The method comprises contacting a population of cells in culture with one or more compounds, and contacting the population of cells with a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane. The detectable signal is measured and the average membrane potential of said population of cells is calculated. The average membrane potential of said population of cells cultured in the presence of the compounds is compared to that of a control population of cells cultured in the absence of the compounds. Compounds that increase the average membrane

potential in said population of cells are identified as candidate compounds for promoting dedifferentiation of cells.

[0027] In a ninth aspect, the invention provides a method of screening for compounds that promote dedifferentiation of cells. The method comprises contacting an animal or tissue with one or more compounds, and contacting said animal or tissue with a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane. The detectable signal is measured and the average membrane potential of one or more populations of cells in the animal or tissue is calculated. The average membrane potential of said population of cells cultured in the presence of the compounds is compared to that of a control population of cells cultured in the absence of the compounds. Compounds that increase the average membrane potential in said population of cells are identified as candidate compounds for promoting dedifferentiation of cells.

[0028] In a tenth aspect, the invention provides a method of screening for compounds that promote dedifferentiation of cells. The method comprises producing an injury in an animal or tissue, contacting said injured animal or tissue with one or more compounds, and contacting said injured animal or tissue with a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane. The detectable signal is measured and the average membrane potential of one or more populations of cells proximate to the injury in the animal or tissue is calculated. The average membrane potential of said population of cells cultured in the presence of the compounds is compared to that of a control population of cells cultured in the absence of the compounds. Compounds that increase the average membrane potential in said one or more populations of cells proximate to the injury are identified as candidate compounds for promoting dedifferentiation of cells.

[0029] In one embodiment of any of the foregoing, the one or more compounds are independently selected from nucleic acids, peptides, proteins, small organic molecules, small inorganic molecules, antisense oligonucleotides, RNAi constructs, and antibodies.

[0030] In an eleventh aspect, the invention provides a method for promoting dedifferentiation. The method comprises providing a population of cells. The cells are contacted with a compound that modulates ion flux mediated by a class of ion transporter proteins. The compound that modulates ion flux mediated by a class of ion transporter proteins promotes depolarization of cell membranes of one or more cells in the population of cells, thereby promoting dedifferentiation of one or more cells in the population of cells.

[0031] In one embodiment, the method further comprises culturing the population of cells which includes one or more dedifferentiated cells, wherein said culturing promotes cellular regeneration.

[0032] In a twelfth aspect, the invention provides a method for inhibiting dedifferentiation. The method comprises providing a population of cells. The cells are contacted with a compound that modulates ion flux mediated by a class of ion transporter proteins. The compound that modulates ion flux mediated by a class of ion transporter proteins inhibits depolarization of cell membranes of one or more cells in the population of cells, thereby inhibiting dedifferentiation of one or more cells in the population of cells.

[0033] In one embodiment, the method is an in vitro method and the population of cells is in culture. In another embodiment, the method is an in vitro method and the popu-

lation of cells is in a preparation of tissue in culture. In another embodiment, the method is an in vivo method and the population of cells is resident in an animal. In one embodiment, when the population of cells is resident in an animal, the animal includes an injury. In another embodiment, when the population of cells is resident in an animal, the animal is a fragment of an animal.

[0034] The above methods can be used in cells that are resident in or derived from any species or organism. Exemplary organisms are animals, although the method can similarly be used to access the role of ion flux in plants, bacteria and other prokaryotes, and fungi. In one embodiment, the cells are derived from or resident in an animal selected from a flatworm, an amphibian, a fish, a reptile, a bird, or a mammal. In one embodiment, the animal is a flatworm and the flatworm is a planarian of the class Turbellaria. In another embodiment, the animal is an amphibian and the amphibian is *Xenopus laevis* or *Xenopus tropicalis*. In yet another embodiment, the animal is a mammal selected from a mouse, rat, cat, dog, rabbit, goat, hamster, pig, sheep, non-human primate, or primate.

[0035] In any of the foregoing, the invention contemplates that the cells or animals, regardless of species, can be at any developmental stage. In one embodiment, the population of cells comprises embryonic, fetal, larval, juvenile, or adult cells. In another embodiment, the population of cells is resident in animal, and the animal is an embryonic, fetal, larval, juvenile, or adult stage animal. In another embodiment, the population of cells comprises fertilized or unfertilized oocytes.

[0036] In one embodiment of any of the foregoing, the compound inhibits ion flux mediated by the class of transporter proteins. In another embodiment, the compound promotes ion flux mediated by the class of transporter proteins.

[0037] In one embodiment of any of the foregoing, the compound is selected from a nucleic acid, a peptide, a protein, a small organic molecule, a small inorganic molecule, an antisense oligonucleotide, an RNAi construct, or an antibody.

[0038] In a thirteenth aspect, the invention provides pharmaceutical preparations of one or more compounds identified by the methods of the present invention.

[0039] In a fourteenth aspect, the invention provides pharmaceutical preparations for promoting dedifferentiation and/or regeneration in one or more cells in a population of cells.

[0040] In a fifteenth aspect, the invention provides pharmaceutical preparations for inhibiting dedifferentiation and/or regeneration in one or more cells in a population of cells.

[0041] In a sixteenth aspect, the invention provides use of a compound that modulates ion flux and/or membrane potential in the manufacture of a medicament for promoting dedifferentiation and/or regeneration.

[0042] In one embodiment, the compound inhibits ion flux mediated by a class of ion transporter proteins, thereby modulating ion flux and/or membrane potential. In another embodiment, the compound promotes ion flux mediated by a class of ion transporter proteins, thereby modulating ion flux and/or membrane potential.

[0043] In a seventeenth aspect, the invention provides use of a compound that modulates ion flux and/or membrane potential in the manufacture of a medicament for inhibiting dedifferentiation and/or regeneration.

[0044] In one embodiment, the compound inhibits ion flux mediated by a class of ion transporter proteins, thereby modulating ion flux and/or membrane potential. In another embodi-

ment, the compound promotes ion flux mediated by a class of ion transporter proteins, thereby modulating ion flux and/or membrane potential.

[0045] In an eighteenth aspect, the invention provides a method for identifying progenitor cells. This aspect of the invention is based on the appreciation of a correlation between sternness and membrane potential. Given this correlation, methods of identifying depolarized cells can be used to identify and/or separate progenitor cells from amongst a population of cells, thereby facilitating further culture, purification, and analysis of progenitor cells. In one embodiment, a method for identifying progenitor cells comprises contacting a population of cells with a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane. One or more cells in the population of cells which have a depolarized cell membrane with a membrane potential of greater than or equal to -20 mV are identified, thereby identifying candidate progenitor cells.

[0046] In a nineteenth aspect, a method for identifying progenitor cells comprises contacting a population of cells with a pH sensitive agent that produces a detectable signal. One or more cells in the population of cells which have an intracellular pH of less than or equal to 6.7 are identified, thereby identifying candidate progenitor cells.

[0047] In any of the foregoing, the method is an in vitro method and the population of cells is in culture. In another embodiment, the method is an in vitro method and the population of cells is in a preparation of tissue in culture. In another embodiment, the method is an in vivo method and the population of cells is resident in an animal.

[0048] The above methods can be used in cells that are resident in or derived from any species or organism. Exemplary organisms are animals, although the method can similarly be used to access the role of ion flux in plants, bacteria and other prokaryotes, and fungi. In one embodiment, the cells are derived from or resident in an animal selected from a flatworm, an amphibian, a fish, a reptile, a bird, or a mammal. In one embodiment, the animal is a flatworm and the flatworm is a planarian of the class Turbellaria. In another embodiment, the animal is an amphibian and the amphibian is *Xenopus laevis* or *Xenopus tropicalis*. In yet another embodiment, the animal is a mammal selected from a mouse, rat, cat, dog, rabbit, goat, hamster, pig, sheep, non-human primate, or primate.

[0049] In any of the foregoing, the invention contemplates that the cells or animals, regardless of species, can be at any developmental stage. In one embodiment, the population of cells comprises embryonic, fetal, larval, juvenile, or adult cells. In another embodiment, the population of cells is resident in animal, and the animal is an embryonic, fetal, larval, juvenile, or adult stage animal. In another embodiment, the population of cells comprises fertilized or unfertilized oocytes.

[0050] In one embodiment, the method comprises identifying one or more cells having a membrane potential greater than or equal to -20 mV and less than or equal to 30 mV. In another embodiment, the method comprises identifying one or more cells having a membrane potential of greater than or equal to -15 mV. In another embodiment, the method comprises identifying one or more cells having a membrane potential of greater than or equal to -10 mV. In still another embodiment, the method comprises identifying one or more cells having a membrane potential of greater than or equal to -5 mV. In yet another embodiment, the method comprises

identifying one or more cells having a membrane potential of greater than or equal to 0 mV, 5 mV, 10 mV, 15 mV, or 20 mV.

[0051] In another embodiment, the method comprises contacting the population of cells with both a pH sensitive agent and a voltage sensitive agent. One or more cells having both an intracellular pH of less than or equal to 6.7 and a membrane potential of greater than or equal to -20 mV are identified, thereby identifying candidate progenitor cells.

[0052] In one embodiment of any of the foregoing, the method may further comprise separating candidate progenitor cells (e.g., all of the identified candidate progenitor cells or a subset of the identified candidate progenitor cells) from the population of the cells. When progenitor cells are separated from the population of cells, the separated progenitor cells may be cultured to produce a population of cells enriched in progenitor cells.

[0053] In one embodiment of any of the foregoing, the detectable signal produced by the agents (e.g., the voltage sensitive agent or the pH sensitive agent) is a fluorescent signal. Exemplary voltage sensitive agents include, but are not limited to, bis-(1,3-dibutylbarbituric acid)pentamethine oxonol (DiBAC₄(5)); bis-(1,3-dibutylbarbituric acid)trimethine oxonol (DiBAC₄(3)); bis-(1,3-diethylthiobarbituric acid)trimethine oxonol (DiSBAC₂(3)); 3,3'-diethyloxacarbocyanine iodide (DiOC₂(3)); 3,3'-diheptyloxacarbocyanine iodide (DiOC₇(3)); 3,3'-dihexyloxacarbocyanine iodide (DiOC₆(3)); 3,3'-dipentyloxacarbocyanine iodide (DiOC₅(3)); 1,1',3,3,3',3'-hexamethylindodicarbocyanine iodide (DiIC₁(5)); a structural variant thereof; or a functional variant thereof. Exemplary pH sensitive agents include, but are not limited to, 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein (BCECF); 5-(and-6)-carboxy SNARF®-1; LysoTracker® Blue DND-22; LysoTracker® Green DND-26; LysoTracker® Red DND-99; LysoTracker® Yellow HCK-123; a structural variant thereof; or a functional variant thereof.

[0054] In certain embodiment, the method comprises contacting cells with both a voltage sensitive agent and a pH sensitive agent. In one embodiment, the cells are contacted with the two agents simultaneously. In another embodiment, the cells are contacted with the two agents sequentially.

[0055] In a related embodiment of the foregoing aspects and embodiments of the invention, the invention provides a method for separating progenitor cells from an animal or tissue. The method comprises contacting an animal or tissue with a voltage sensitive agent that produces a detectable signal in cells in response to a depolarized cell membrane. One or more cells in the animal or tissue having a depolarized membrane potential of greater than or equal to -20 mV are identified, thereby identifying candidate progenitor cells. The identified progenitor cells are then removed from the animal or tissue, thereby separating the progenitor cells from the remainder of the animal or tissue.

[0056] In another related embodiment of the foregoing aspects and embodiments of the invention, the invention provides a method for separating progenitor cells from an animal or tissue. The method comprises contacting an animal or tissue with a pH sensitive agent that produces a detectable signal. One or more cells in the animal or tissue having an intracellular pH of less than or equal to 6.7 are identified, thereby identifying candidate progenitor cells. The identified progenitor cells are then removed from the animal or tissue, thereby separating the progenitor cells from the remainder of the animal or tissue.

[0057] In yet another related embodiment, the invention provides a method for separating progenitor cells from an animal or tissue. The method comprises contacting an animal or tissue with both a pH sensitive agent that produces a detectable signal and a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane. One or more cells in the animal or tissue having both a depolarized membrane potential of greater than or equal to -20 mV and an intracellular pH of less than or equal to 6.7 are identified, thereby identifying candidate progenitor cells. The identified progenitor cells are then removed from the animal or tissue, thereby separating the progenitor cells from the remainder of the animal or tissue.

[0058] In one embodiment removing identified progenitor cells comprises dissecting out candidate progenitor cells, thereby removing candidate progenitor cells from the animal or tissue. In another embodiment, removing identified candidate progenitor cells comprises dissociating the animal or tissue; and sorting candidate progenitor cells, thereby separating candidate progenitor cells from the animal or tissue. When progenitor cells are sorted, the method of sorting can comprise sorting based on the detectable signal (e.g., the detectable signal provided by the pH sensitive agent or the voltage sensitive agent). An exemplary method of cell sorting based on a detectable signal fluorescence activated cell sorting (FACS) analysis.

[0059] The present invention also provides methods for identifying whether ion flux is involved in a particular biological process, and if so, which class of ion transporter proteins may mediate that biological process. The present invention further provides a variety of methods based on the role of ion transporter proteins and ion flux during cellular dedifferentiation and/or regeneration.

[0060] In a twentieth aspect, the invention provides a method for determining whether ion flux is involved in a particular biological process, and if so, identifying a class of ion transporter proteins that mediate ion flux during the particular biological process. A method for identifying a class of ion transporter proteins which mediate ion flux during a particular biological process comprises providing a population of cells that can be used to measure a particular biological process. The population of cells is contacted with a compound that modulates ion flux mediated by a class of ion transporter proteins. Following administration of a compound that modulates ion flux, the particular biological process is measured or otherwise assayed, and compared in the presence versus the absence of the compound. If there is a change in the particular biological process in the presence versus the absence of the compound (e.g., the compound that modulates ion flux mediated by a class of ion transporters), then that class of ion transporter proteins is identified as a candidate for mediating (in whole or in part) ion flux during that particular biological process in the population of cells.

[0061] In a twenty first aspect, the invention provides a method for identifying a class of ion transporter proteins that mediate ion flux during a particular biological process. The method is a reiterative method comprising assessing the effects of compounds that are increasingly specific. In other words, in each successive round of screening, the cells are contacted with a compound that modulates the activity of an increasingly specific/defined class of ion transporter proteins. By way of example, a method for identifying a class of ion transporter proteins which mediate ion flux during a particular biological process comprises providing a population of

cells that can be used to measure a particular biological process. The population of cells is contacted with a first compound that modulates ion flux mediated by a first class of ion transporter proteins. Following administration of the first compound that modulates ion flux, the particular biological process is measured or otherwise assayed, and compared in the presence versus the absence of the compound. If there is a change in the particular biological process in the presence versus the absence of the compound (e.g., the compound that modulates ion flux mediated by a class of ion transporters), then that class of ion transporter proteins is identified as a candidate for mediating (in whole or in part) ion flux during that particular biological process in the population of cells. A second population of the equivalent cells is then contacted with a second compound that modulates ion flux mediated by a second class of ion transporter proteins. This second class of ion transporter proteins comprises a subset of the first class of ion transporter proteins, and thus the second compound serves to further narrow/specify the candidate class of ion transporter proteins that mediate ion flux and thus mediate the particular biological process. Following contacting the population of cells with the second compound, the method comprises measuring the particular biological process in the population of cells in the presence of the second compound versus the absence of the second compound, and determining whether the second compound that modulates ion flux mediated by the second class of ion transporter proteins changes the particular biological process in the population of cells. This method facilitates identification of a candidate class of ion transporter proteins that is more specific than that identified following the use of a single round of screening with a single compound.

[0062] In any of the foregoing, the method may comprise a reiterative method where the screening steps are repeated multiple times. In certain embodiments, a compound that modulates an increasingly specific class of ion transporter proteins is used in each successive round of screening. In this way, each successive round of screening serves to further define and narrow the candidate class of ion transporter proteins that may play a role in the particular biological process being studies. In this way, the invention provides an ordered, hierarchical screening approach for (i) identifying whether ion transport is involved in a particular biological process, (ii) identifying, through one or more rounds of successive screening, one or more candidate classes of ion transporter proteins that may mediate ion transport in a particular biological process, and in certain embodiments (iii) identifying one or more members of a particular class of ion transporter proteins capable of mediating ion flux involved in a particular biological process.

[0063] In one embodiment, the method is an in vitro method and the population of cells is in culture. In another embodiment, the method is an in vitro method and the population of cells is in a preparation of tissue in culture. In another embodiment, the method is an in vivo method and the population of cells is resident in an animal.

[0064] The above methods can be used in cells that are resident in or derived from any species or organism. Exemplary organisms are animals, although the method can similarly be used to access the role of ion flux in plants, bacteria and other prokaryotes, and fungi. In one embodiment, the cells are derived from or resident in an animal selected from a flatworm, an amphibian, a fish, a reptile, a bird, or a mammal. In one embodiment, the animal is a flatworm and the

flatworm is a planarian of the class Turbellaria. In another embodiment, the animal is an amphibian and the amphibian is *Xenopus laevis* or *Xenopus tropicalis*. In yet another embodiment, the animal is a mammal selected from a mouse, rat, cat, dog, rabbit, goat, hamster, pig, sheep, non-human primate, or primate.

[0065] In any of the foregoing, the invention contemplates that the cells or animals, regardless of species, can be at any developmental stage. In one embodiment, the population of cells comprises embryonic, fetal, larval, juvenile, or adult cells. In another embodiment, the population of cells is resident in an animal, and the animal is an embryonic, fetal, larval, juvenile, or adult stage animal. In another embodiment, the population of cells comprises fertilized or unfertilized oocytes.

[0066] The foregoing methods comprise contacting cells with a compound that modulates ion flux mediated by the class of ion transporter proteins. In one embodiment, the compound inhibits ion flux mediated by the class of ion transporter proteins. In another embodiment, the compound promotes ion flux mediated by the class of transporter proteins. An inhibitor or promoter of ion flux may, for example, modulate the expression and/or activity of one or more ion transporters within a class of ion transporters. Compounds (e.g., inhibitors or promoters) may act directly upon one or more ion transporter protein (e.g., by directly binding to ion transporter proteins) or compounds may act indirectly (e.g., via a necessary cofactor, by influencing transcription or translation of ion transporter proteins, etc.).

[0067] The foregoing methods comprise contacting cells with a compound that modulates ion flux mediated by a class of ion transporter proteins. In one embodiment, the compound is selected from nucleic acids, peptides, proteins, small organic molecules, small inorganic molecules, antisense oligonucleotides, RNAi constructs, or antibodies. When the methods comprise a reiterative method involving multiple rounds of screening using compounds that are increasing selective with respect to the class of ion transporter proteins, the one or more compounds used are independently selected from nucleic acids, peptides, proteins, small organic molecules, small inorganic molecules, antisense oligonucleotides, RNAi constructs, or antibodies.

[0068] The foregoing methods can be used to study any of a range of biological processes. The biological process can be defined generally, as well as specifically based on the model organism, developmental stage, and cell type under investigation. In one embodiment, the biological process is a general biological process selected from cell proliferation, cell differentiation, apoptosis, cell survival, cell migration, regeneration, or dedifferentiation. In certain embodiments, the particular biological process can be further described based on the cell type, organism, or stage of development being evaluated.

[0069] Numerous methods in cell and developmental biology can be used to assay the particular biological process. In one embodiment, measuring the particular biological process comprises measuring a change in gene expression, a change in protein expression, or a change in morphology. In another embodiment, measuring the particular biological process comprises measuring a change in the rate or extent of cell proliferation, a change in cell differentiation, a change in the rate of tissue regeneration, or a change in the rate or extent of apoptosis.

[0070] The foregoing provide a method for identifying a candidate class of ion transporter proteins that may mediate ion flux, thereby modulating a particular biological process. Identification of a candidate class of ion transporter proteins provides a list of candidate ion transporter proteins that may be involved in a particular biological process. In certain embodiments, the method further comprises evaluating one or more individual ion transporter proteins within the identified class to identify a particular ion transporter protein involved in modulating the particular biological process.

[0071] In one embodiment, evaluating one or more individual ion transporter proteins (e.g., individual proteins that are members of an identified class of transporters) comprises examining gene or protein expression of one or more individual ion transporter proteins.

[0072] In another embodiment, evaluating one or more individual ion transporter proteins comprises providing a population of cells that can be used to measure a particular biological process, and inhibiting expression or activity of one or more individual ion transporter proteins, which ion transporter proteins are members of the candidate class of ion transporter proteins. In another embodiment, evaluating one or more individual ion transporter proteins comprises providing a population of cells that can be used to measure a particular biological process, and promoting expression or activity of one or more individual ion transporter proteins, which ion transporter proteins are members of the candidate class of ion transporter proteins.

[0073] Following inhibiting or promoting expression or activity of one or more individual ion transporter proteins, the particular biological process can be assayed in the cells to determine whether inhibition or promotion of the expression or activity of said one or more ion transporter proteins changes the particular biological process in the population of cells. The invention contemplates a combinatorial strategy in which both expression and function of the particular one or more ion transporter proteins is evaluated.

[0074] In one embodiment of methods used to further identify one or more particular ion transporter proteins, inhibiting the expression or activity of the ion transporter protein may comprise contacting the population of cells with an agent that specifically inhibits the expression or activity of the ion transporter protein and does not substantially inhibit the expression or activity of other ion transporter proteins that are a member of the candidate class of ion transporter proteins. In one embodiment, the agent that inhibits expression or activity of an ion transporter protein is an RNAi construct that specifically inhibits the expression or activity of the ion transporter protein, and the method comprises contacting the population of cells with the RNAi construct. In another embodiment, the agent that inhibits expression or activity of an ion transporter protein is an antibody that is immunoreactive with and specifically inhibits the activity of the ion transporter protein, and the method comprises contacting the population of cells with the antibody.

[0075] In one embodiment, inhibiting the expression or activity of an ion transporter protein comprises contacting the population of cells with any of a nucleic acid, peptide, protein, small organic molecule, small inorganic molecule, antisense oligonucleotide, RNAi construct, or antibody, thereby specifically inhibiting the expression or activity of the ion transporter protein. In one embodiment, the agent is a nucleic acid and the nucleic acid encodes a dominant negative form of

the ion transporter protein. In another embodiment, the agent is a protein, and the protein is a dominant negative form of the ion transporter protein.

[0076] In one embodiment, promoting the expression or activity of an ion transporter protein comprises contacting the population of cells with any of a nucleic acid, peptide, protein, small organic molecule, or small inorganic molecule, thereby specifically inhibiting the expression or activity of the ion transporter protein. In one embodiment, the agent is a nucleic acid and the nucleic acid encodes the candidate ion transporter protein. In another embodiment, the agent is a protein, and the protein is the ion transporter protein. Ectopic expression of the candidate ion transporter protein (e.g., using a nucleic acid or protein corresponding to the candidate ion transporter protein) can be used to promote expression of an ion transporter protein in a population of cells.

[0077] In any of the foregoing embodiments of this aspect of the invention, the invention further contemplates assaying the population of cells to determine whether a compound that modulates ion flux via a class of ion transporter proteins or via a particular candidate ion transporter protein also alters the membrane potential of cells within the population of cells. In one embodiment, the membrane potential is assessed using a vibrating probe. In another embodiment, the membrane potential is assessed using a fluorescent agent that emits a detectable signal indicative of the membrane potential.

[0078] In any of the foregoing embodiments of this aspect of the invention which call for multiple rounds of screening and analysis, the invention contemplates using the same or equivalent population of cells (e.g., populations that are derived from the same species, tissue type, and are of the same stage of developmental). Unless specifically stated, reference to the population of cells is not meant to imply that the identical cells are used in subsequent rounds of study. Rather, the term is meant to indicate that equivalent cells are used when conducting experiments that require multiple rounds of screening. However, the invention contemplates that it may sometimes be advantageous to conduct various rounds of screening using different populations of cells (e.g., populations of cells derived from a different species, tissue type, or stage of development).

[0079] In any of the foregoing embodiments of this aspect of the invention, the invention further contemplates screening methods for other families of proteins susceptible to hierarchical screening. For example, other susceptible families of proteins include neurotransmitters and molecular motors.

[0080] The invention contemplates combinations of any of the foregoing aspects and embodiments of the invention. Furthermore, in any of the foregoing aspects and embodiments of the invention, the invention contemplates the use of compounds to specifically modulate a particular biological process (e.g., to modulate a biological process directly rather than by generally perturbing viability of the organism).

BRIEF DESCRIPTION OF THE DRAWINGS

[0081] The invention may be better understood by referring to the following description of illustrative embodiments, taken in conjunction with the accompanying drawings, in which like reference designations refer to like components and depicted components are not necessarily drawn to scale.

[0082] FIG. 1 illustrates a multi-step approach for evaluating a role for calcium ion flux in a biological process.

[0083] FIG. 2 illustrates a multi-step approach for evaluating a role for potassium ion flux in a biological process.

[0084] FIG. 3 illustrates a multi-step approach for evaluating a role for hydrogen ion flux in a biological process.

[0085] FIG. 4 illustrates a multi-step approach for evaluating a role for sodium ion flux in a biological process.

[0086] FIG. 5 illustrates a multi-step approach for evaluating a role for chloride ion flux in a biological process.

[0087] FIG. 6 illustrates a method for evaluating a role for ion flux in a biological process.

[0088] FIG. 7 illustrates a specific example of a multi-step approach for determining a class of ion transporter proteins that mediate calcium ion flux, and thereby modulate a particular biological process.

[0089] FIGS. 8A-8F illustrate standard curves used to interpret and utilize fluorescent dyes.

[0090] FIG. 9 summarizes the results of a candidate screen to identify ion transporter proteins involved in regeneration in *Xenopus* tails.

[0091] FIG. 10A-10F illustrate the results of manipulation of ion flux on regeneration in *Xenopus* tails.

[0092] FIG. 11A-11F depict membrane voltage and pH in regenerating tails in the presence and absence of compound that modulate ion flux mediated by ion transporter proteins.

[0093] FIG. 12A-12J show that the V-ATPase is required for tail regeneration in *Xenopus*.

[0094] FIG. 13A-13M show the characterization of expression of V-ATPase and physiology in the regeneration bud.

[0095] FIG. 14A-14N show V-ATPase function is required for the up-regulation of cell proliferation and axonal patterning.

[0096] FIG. 15 summarizes a step-wise model of tail regeneration consisting of physiological, gene expression, and morphogenetic modules.

[0097] FIG. 16 illustrates one embodiment of a method of promoting regeneration.

DETAILED DESCRIPTION OF THE INVENTION

(i) Overview

[0098] The present invention is based on our appreciation of the important role of ion flux and membrane potential in a range of biological processes. Ion flows set up by channels, pumps, and other ion transporter proteins produce pH and voltage gradients within cells and across cells fields. Ion flow and the regulation of ion flow plays an important role in a range of biological processes, and thus methods of identifying ion transporter proteins involved in particular biological processes are critical in understanding and manipulating a variety of physiological phenomenon in cells, tissues, and organisms.

[0099] We have uncovered a role for ion flux in a biological process of tremendous interest to scientists and physicians alike: cellular dedifferentiation and regeneration. In light of the role of ion flux in dedifferentiation and regeneration, compounds that modulate ion flux may be used to promote dedifferentiation and/or regeneration, thereby promoting regeneration of cells and tissues in vitro or in vivo. Furthermore, compounds that modulate ion flux may be used to inhibit dedifferentiation and/or regeneration, thereby inhibiting regeneration of cells and tissues in vitro or in vivo. For any of the foregoing, these compounds may be compounds already known to influence ion flux and membrane potential. Known modulators of ion flux and/or of the expression or activity of ion transporter proteins can be tested in a particular cell type or system to identify which of the known compounds

modulate dedifferentiation and/or regeneration. Alternatively, compounds previously unknown to influence ion flux and membrane potential can be tested to identify compounds that both influence ion flux and membrane potential and modulate dedifferentiation and/or regeneration. In yet another alternative, novel compounds can be generated and tested for the ability to (i) modulate ion flux and membrane potential (e.g., by influencing the expression or activity of one or more ion transporter proteins) and (ii) modulate dedifferentiation and/or regeneration in one or more model systems.

[0100] The present invention also provides additional methods based on an appreciation for the importance of ion flux in a range of biological process. In one aspect, the invention provides methods for identifying and/or isolating progenitor cells based on membrane potential or pH characteristics typically associated with progenitor cells. In another aspect, the invention provides a hierarchical screening approach that allows ordered identification and characterization for whether a particular biological process is mediated, in whole or in part, by modulation of membrane potential. This ordered approach is unique to previous candidate screening approaches and permits efficient identification of (i) whether a particular process is mediated by modulation of membrane potential and (ii) what class or subclass of ion transporter protein mediates membrane potential in that particular system or process.

[0101] Finally, the present invention contemplates that the various screening methods of the present invention can be used in combination as part of an ordered approach for studying a process in a given biological system (e.g., regeneration) and selecting appropriate agents that modulate ion flux in that system. Use of a combination of the methods described herein to study a biological process is depicted in FIG. 16.

(ii) Definitions

[0102] The articles “a” and “an” are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

[0103] As used herein, “protein” is a polymer consisting essentially of any of the 20 amino acids. Although “polypeptide” is often used in reference to relatively large polypeptides, and “peptide” is often used in reference to small polypeptides, usage of these terms in the art overlaps and is varied.

[0104] The term “wild type” refers to the naturally-occurring polynucleotide sequence encoding a protein, or a portion thereof, or protein sequence, or portion thereof, respectively, as it normally exists *in vivo*. The term “wild type” also refers to a phenotypically and genotypically normal organism.

[0105] The term “mutant” refers to any change in the genetic material of an organism, in particular a change (i.e., deletion, substitution, addition, or alteration) in a wildtype polynucleotide sequence or any change in a wildtype protein sequence. The term “variant” is used interchangeably with “mutant”. Although it is often assumed that a change in the genetic material results in a change of the function of the protein, the terms “mutant” and “variant” refer to a change in the sequence of a wildtype protein regardless of whether that change alters the function of the protein (e.g., increases, decreases, imparts a new function), or whether that change has no effect on the function of the protein (e.g., the mutation or variation is silent). The term “mutant” also refers to an

organism with one or more phenotypic or genotypic alterations in comparison to a wild type organism of the same species.

[0106] As used herein, the term “nucleic acid” refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides.

[0107] As used herein, the term “gene” or “recombinant gene” refers to a nucleic acid comprising an open reading frame encoding a polypeptide, including both exon and (optionally) intron sequences.

[0108] As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as “expression vectors”.

[0109] A polynucleotide sequence (DNA, RNA) is “operatively linked” to an expression control sequence when the expression control sequence controls and regulates the transcription and translation of that polynucleotide sequence. The term “operatively linked” includes having an appropriate start signal (e.g., ATG) in front of the polynucleotide sequence to be expressed, and maintaining the correct reading frame to permit expression of the polynucleotide sequence under the control of the expression control sequence, and production of the desired polypeptide encoded by the polynucleotide sequence.

[0110] The terms “compound” and “agent” are used interchangeably to refer to nucleic acids, peptides, polypeptides, or small molecules. In the context of the present invention, compounds or agents may modulate ion flux, for example, by inhibiting or promoting ion flux mediated by a particular ion transporter protein or class of ion transporter proteins. Exemplary nucleic acid agents include, but are not limited to, sense or antisense nucleic acids, sense or antisense oligonucleotides, ribozymes, and RNAi constructs. Exemplary peptide and polypeptide agents include growth factors, transcription factors, peptidomimetics, and antibodies, as well as particular ion transporter proteins or subunits thereof. Exemplary small molecules include small organic or inorganic molecules, e.g., with molecular weights less than 7500 amu, preferably less than 5000 amu, and even more preferably less than 2000, 1500, 1000, or 500 amu. One class of small organic or inorganic molecules is non-peptidyl, e.g., containing 2, 1, or no peptide and/or saccharide linkages. The term agent is also used, when specified, to refer to compounds that produce a detectable signal in response to pH or membrane voltage. In this context, the agent serves as an indicator of intracellular pH or membrane voltage. Exemplary agents include fluorescent dyes that provide a detectable signal indicative of membrane voltage or pH.

[0111] For purposes of this invention, the chemical elements are identified in accordance with the Periodic Table of the Elements, CAS version, Handbook of Chemistry and Physics, 67th Ed., 1986-87, inside cover.

[0112] A “marker” is used to determine the state of a cell. Markers are characteristics, whether morphological or biochemical (enzymatic), particular to a cell type, or molecules expressed by the cell type. A marker may be a protein marker,

such as a protein marker possessing an epitope for antibodies or other binding molecules available in the art. A marker may also consist of any molecule found in a cell, including, but not limited to, proteins (peptides and polypeptides), lipids, polysaccharides, nucleic acids and steroids. Additionally, a marker may comprise a morphological or functional characteristic of a cell. Examples of morphological traits include, but are not limited to, shape, size, and nuclear to cytoplasmic ratio. Examples of functional traits include, but are not limited to, the ability to adhere to particular substrates, ability to incorporate or exclude particular dyes, ability to migrate under particular conditions, and the ability to differentiate along particular lineages.

[0113] Markers may be detected by any method available to one of skill in the art. In addition to antibodies (and all antibody derivatives) that recognize and bind at least one epitope on a marker molecule, markers may be detected using analytical techniques, such as by protein dot blots, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), or any other gel system that separates proteins, with subsequent visualization of the marker (such as Western blots), gel filtration, affinity column purification; morphologically, such as fluorescent-activated cell sorting (FACS), staining with dyes that have a specific reaction with a marker molecule (such as ruthenium red and extracellular matrix molecules), specific morphological characteristics (such as the presence of microvilli in epithelia, or the pseudopodia/filopodia in migrating cells, such as fibroblasts and mesenchyme); and biochemically, such as assaying for an enzymatic product or intermediate, or the overall composition of a cell, such as the ratio of protein to lipid, or lipid to sugar, or even the ratio of two specific lipids to each other, or polysaccharides. In the case of nucleic acid markers, any known method may be used. If such a marker is a nucleic acid, PCR, RT-PCR, in situ hybridization, dot blot hybridization, Northern blots, Southern blots and the like may be used, coupled with suitable detection methods. If such a marker is a morphological and/or functional trait, suitable methods include visual inspection using, for example, the unaided eye, a stereomicroscope, a dissecting microscope, a confocal microscope, or an electron microscope.

[0114] "Differentiation" describes the acquisition or possession of one or more characteristics or functions different from that of the original cell type. A differentiated cell is one that has a different character or function from the surrounding structures or from the precursor of that cell (even the same cell). The process of differentiation gives rise from a limited set of cells (for example, in vertebrates, the three germ layers of the embryo: ectoderm, mesoderm and endoderm) to cellular diversity, creating all of the many specialized cell types that comprise an individual.

[0115] Differentiation is a developmental process whereby cells assume a specialized phenotype, e.g., acquire one or more characteristics or functions distinct from other cell types. In some cases, the differentiated phenotype refers to a cell phenotype that is at the mature endpoint in some developmental pathway. In many, but not all tissues, the process of differentiation is coupled with exit from the cell cycle. In these cases, the cells typically lose or greatly restrict their capacity to proliferate and such cells are commonly referred to as being terminally differentiated.

[0116] The term regeneration refers to the restoration of cells, tissues, or structures following injury, ablation, loss, or disease. Regeneration involves an interplay of proliferation,

differentiation, sometimes dedifferentiation. In some instances, regeneration refers to individual cells or groups of cells. In other instances, regeneration comprises restoration of all or a portion of a tissue or organ. The invention provides methods of promoting or enhancing regeneration. In some embodiments, the method of promoting or enhancing regeneration includes modulating one or more of proliferation, differentiation, dedifferentiation, survival, or migration.

[0117] As used herein, the term "non-regenerating cells" refers to non-naturally regenerating cells from a non-regenerating organism or cells from a regenerating organism in a refractory period. The term "naturally regenerating cells" refers to cells from a regenerating organism that are typically capable of regenerating. In certain embodiments naturally regenerating cells may be in a regenerating state.

[0118] As used herein, the term "population of cells" refers to one or more cells in a tissue or organ. A tissue or organ of the invention may be part of an organism or cultured in vitro.

[0119] As used herein, the term "effective amount" means the total amount of the active component(s) of a composition or compound that is sufficient to cause a statistically significant change on a detectable biochemical or phenotypic characteristic. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the effect, whether administered in combination, serially or simultaneously.

[0120] The term "membrane" refers to phospholipid bilayers.

[0121] The term "ion flux" refers to the movement of ions through an area/unit time. The term does not imply anything about the mechanism of ion movement. The term includes ion flux mediated by any ion transporter protein regardless of whether the transporter protein actively or passively shuttles ions. The term ion flux includes movement of ions into a cell or movement of ions out of a cell (e.g., efflux or influx).

[0122] As used herein, the terms "ion transporter proteins" and "transporter proteins" are used interchangeably and include proteins that mediate ion flux regardless of the particular ion species transported or the particular mechanism of action. The term includes proteins that are passive transporters, as well as proteins that are active transporters. "Class of ion transporter proteins" refers to categories of transporter proteins organized based on similar functional characteristics. For examples, a class of ion transporter proteins may include transporter proteins that transport a particular ion species (e.g., Ca, Na, H) or transporter proteins that transport a particular ion species using a particular mechanism of action.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

(i) Detailed Description of Methods and Apparatuses

[0123] (a) Methods of Modulating Dedifferentiation and/or Regeneration

[0124] The regeneration of complex tissues and organ systems lost to injury, senescence, or disease is a key goal of biomedicine. In addition to its clinical applications, the regeneration of organs is fascinating because it represents one of the most fundamental properties of most living things: recognition of damage, and self-repair. Currently, substantial efforts are being invested in academia and industry to under-

stand the fundamental principles influencing dedifferentiation and regeneration so that the powers of these processes can be harnessed and used to treat degenerative diseases and injuries.

[0125] Animal regeneration can be conceptually divided into four phases: 1) injury processes which initiate regeneration, 2) formation of a blastema including in some cases the de-differentiation of local cells, 3) the differentiation of blastema components, and 4) morphogenesis of appropriate structures. Additionally, it has long been recognized that certain species appear to have a more robust regenerative capacity (e.g., planaria, amphibians, fish) while other species appear to have a less robust regenerative capacity. Thus, one goal of scientists studying regeneration has been to understand the principles and processes that modulate de-differentiation and/or regeneration in species capable of mounting robust regenerative responses, so that these principles and processes can be applied to increase the regenerative capacity of other organisms.

[0126] This aspect of the present invention is based, in part, on our findings indicating that modulation of membrane potential and/or pH can be used to reset differentiated cells. Without being bound by theory, the resetting of cells may promote their dedifferentiation. Once dedifferentiated, these cells are responsive to endogenous or exogenously supplied cues that drive proliferation and other processes. As a result, regeneration of tissues can occur. Accordingly, by modulating membrane potential and/or pH, cells, tissues, and organs can be regenerated. As outlined further below, studies performed in organisms and tissues that naturally regenerate indicates that there is a window of membrane potential and/or pH permissive for regeneration. Thus, by modulating membrane potential and/or pH to within this regeneration permissive range, the present invention provides methods for promoting regeneration in any of a variety of species—including species and tissues that do not have a naturally high regenerative potential.

[0127] In certain aspects, the present invention provides a method of promoting regeneration by modifying the membrane potential and/or pH of one or more cells. Promoting regeneration in non-regenerating cells may be accomplished by first determining the membrane potential and/or pH range permissive for regeneration in naturally regenerating cells. This membrane potential and/or pH range may be determined experimentally or by referring to known values. The membrane potential and/or pH range is determined in the population of non-regenerating cells of interest. Again, this membrane potential and/or pH range may be determined experimentally or by referring to known values. In certain embodiments, the permissive membrane potential and/or pH range will be the same in naturally regenerating cells and non-regenerating cells and will not need to be determined for each tissue or organ or organism. The population of non-regenerating cells is contacted with an agent that modulates ion flux mediated by a class of ion transporter proteins. The agent modifies the membrane potential and/or pH of one or more cells in the population of non-regenerating cells to the range permissive for regeneration as determined experimentally or by referring to known values; thereby promoting regeneration of one or more cells in the population of cells. Accordingly, compounds that modulate ion flux (e.g., inhibitor/promoters of the expression or activity of an ion transporter protein or a class of ion transporter proteins) can be used in methods for promoting dedifferentiation and/or

regeneration. Thus, the present invention provides methods, compositions, and pharmaceutical compositions that can be used to promote cell dedifferentiation and/or regeneration. The methods and compositions of the invention can be used to increase our understanding of the principles underlying dedifferentiation and regeneration. The methods and compositions of the invention can be used to promote dedifferentiation and regeneration *in vivo* or *in vitro*. The methods and compositions of the invention can be used as a basis for the development of therapeutic methods for treating degenerative diseases and injuries that could be ameliorated using methods that enhance regenerative capacity.

[0128] In certain aspects, promotion of regeneration occurs through a biological program. In other words, modulation of membrane potential or pH resets cells to a state permissive for regeneration. The reset cells are activated to carryout the endogenously present developmental pathways that lead to coordinated proliferation, differentiation, and migration of the multiple cell types that constitute complex tissues and organs. Without being bound by theory and in this embodiment of the invention, once the membrane potential of cells is modulated into the regeneration permissive range and the cells are developmentally reset, further signaling cascades and proteins needed to promote coordinated differentiation need not be exogenously supplied, but rather, the endogenous programs that exist during normal development and during regeneration in regenerating organisms are activated within the reset cells or populations of cells.

[0129] In certain aspects, modulation of ion flux in a small number of cells is sufficient to promote dedifferentiation and/or regeneration of an entire tissue or organ. In certain embodiments, the cells that are modulated to promote dedifferentiation and/or regeneration are progenitor cells. In certain embodiments, the cells that are modulated to promote dedifferentiation and/or regeneration are a single cell type. In certain embodiments, the cell type or types that are modulated in order to promote dedifferentiation and/or regeneration are not specifically determined.

[0130] Illustrative examples whereby ion flux can be manipulated to promote dedifferentiation and/or regeneration are provided in the examples. Briefly, compounds that specifically modulate (inhibit or promote) ion flux, membrane potential, and/or pH mediated by a specific ion transporter protein or a class of ion transporter proteins can be administered to cells, tissues, or organisms. In one embodiment, the method comprising modulating regeneration in a system with enhanced regenerative capacity (e.g., planaria, *Xenopus* tail, *Xenopus* limb, zebrafish tail). In another embodiment, the method comprises promoting regeneration (e.g., enhancing regenerative capacity) in a system that does not endogenously have a robust regenerative capacity (FIG. 16).

[0131] Compounds identified as specifically modulating ion flux during regeneration can be formulated and administered as part of a method for promoting dedifferentiation and/or regeneration *in vitro* or *in vivo*. Such compounds can be administered as a composition or a pharmaceutical preparation. Such compounds can be developed and used as part of a therapeutic regimen.

[0132] In certain embodiments, any ion channel or pump may be expressed or modulated to modulate membrane potential, thereby promoting the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, the ion channel or pump may come from any organism and

may originate from a different organism than the one being treated. Some representative channels and pumps are listed in Table 5 (the nucleic acid and amino acid sequences disclosed at the representative accession numbers are hereby incorporated by reference).

[0133] The present invention provide methods for identifying membrane potential and/or pH ranges permissive for regeneration, as well as methods for modulating membrane potential or pH into a regeneration permissive range. Based on these methods, the regenerative ability in tissues, organs, and organisms that are naturally refractory to regeneration can be enhanced.

[0134] In certain embodiments, membrane potential or pH can be modulated using any transporter protein (e.g., channel or pump) capable of modulating membrane potential into the permissive range (e.g., either hyperpolarizing or depolarizing membrane potential depending on the membrane potential of a give cell relative to the permissive range). Without being bound by theory, in such embodiments, the membrane potential itself, rather than the particular ion(s) being transporter, is critical for promoting regeneration. In other embodiments, membrane potential can be modulated using specific ion transporter proteins or classes of ion transporter proteins (e.g., channel or pump). Without being bound by theory, in such embodiments, both the membrane potential itself and the particular ion(s) being transporter may be important for regeneration.

[0135] In certain embodiments, the V-ATPase H⁺ pump, as well as compounds that promote the activity of the V-ATPase H⁺ pump, may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, the P-type H⁺ ATPase, as well as compounds that promote the activity of the P-type H⁺ ATPase, may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, K⁺ channels, as well as compounds that promote the activity of K⁺ channels, may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, the ROMK K⁺ channel may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, the ERG K⁺ channel may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, voltage-gated Na⁺ channels, as well as compounds that promote the activity of voltage-gated Na⁺ channels, may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue. In certain embodiments, more than one pump or channel may be expressed to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue.

[0136] Note that particular cell and tissue types for which promoting regeneration is desirable may vary with respect to their membrane potential and pH characteristics. These characteristics may vary based on species, cell type, age, health, and other factors. Accordingly, modulating membrane potential or pH into a regeneration permissive range sometimes comprises hyperpolarizing the cell membrane and sometimes

comprises depolarizing the cell membrane. However, given the extensive knowledge of ion transporters (e.g., hyperpolarizing and depolarizing ion transporters) and the existence of inhibitors and agonists of ion transporter function, one of skill in the art can readily select agents that hyperpolarize or depolarize—depending on the membrane potential relative to the permissive range.

[0137] Further embodiments of the foregoing methods are described in sections ii-v. Combinations of any of the foregoing and following embodiments are contemplated. Additionally, further discussions of ion transporter proteins and agents that modulate membrane potential are found throughout the specification.

[0138] (b) Ion Transporters in Biological Processes: An Efficient Screening Method

[0139] To generate an electric field, cell membranes provide a power source and generate a voltage potential by segregating ions across the barrier (resulting in membrane voltage). Because ions are charged, they cannot cross membranes by themselves. By establishing membrane bound compartments, cells use the movement of ions to establish ion concentration differences or gradients. Cells use stored energy in the form of these gradients (by carefully regulating the flow of ions across membranes) to accomplish everything, either by coupling local ion flux directly to local physiology (e.g. Ca²⁺ regulation of secretion) or by siphoning energy from the gradient and storing it as ATP.

[0140] Generally, the myriad of mechanisms by which ion transporter proteins shuttle ions and establish membrane potential are described using four characteristics: (a) the ion or ions that move; (b) the number of ions moved, and the relative direction of the ions when a transporter moves multiple ions; (c) whether the ion transporter protein undergoes a conformational change during transport; and (d) whether any of the ions move up its concentration gradient. This fourth characteristic determines whether movement of the ion requires energy (e.g., generated by ATP hydrolysis, generated via coupling the energy requiring process of moving up a gradient to the energy releasing process of moving down a gradient).

[0141] The first characteristic for describing an ion transporter protein is according to the particular ion or ions it translocates. The following ions generally move through biological membranes: H⁺, Cl⁻, Na⁺, K⁺, and Ca²⁺, and to a lesser degree, Fe²⁺, Cu²⁺, and Zn²⁺. Additionally, charged species such as HCO₃⁻ and OH⁻ can move via ion transporter proteins across cell membranes.

[0142] Most transporter proteins such as channels and pumps are highly specific with respect to the ion species translocated. However, gap junctions are ion transporter proteins that permit non-specific (with respect to ion species) ion flux (Bruzzone et al., 1996, Bioessays 18: 709-718; Nicholson, 2003, J Cell Sci 116: 4479-4481).

[0143] The second characteristic is the number of ion species shuttled by a particular transporter. When a transporter translocates a single ion species (e.g., only Cl⁻ ions or only Ca²⁺ ions), it is often referred to as a uniporter. When a transporter translocates two or more ion species and shuttles the two or more species in the same direction, the transporter is often referred to as a symporter or a cotransporter. When a transporter translocates two or more ion species but shuttles species in opposite directions relative to one another, the transporter is often referred to as an antiport or an exchanger.

[0144] The third characteristic deals with whether the transporter undergoes a conformational change during transit of the ion species. Ion transporter proteins that remain open at both ends during ion translocation (e.g., at both their intracellular and extracellular end), are often referred to as channels. Ion transporter proteins that undergo a conformational change during translocation so that one end of the transporter open at a time are often referred to as transporters. Ion transporter proteins that physically travel from one side of the membrane to the other during translocation of an ion species are often referred to as either transporters, carriers, or ionophores.

[0145] We note that ion transporter proteins can also be described based on whether the ends of the protein are opened or closed in the absence of active ion translocation. Proteins that remain closed (e.g., the intra and extracellular openings are closed) in the absence of ion translocation are often termed gated proteins. The foregoing terminology is often used in the art to help readily categorize various transporter proteins. Throughout the application however, unless reference to a particular protein is being made, we will use the term "ion transporter protein" generically to refer to any transporter protein regardless of mechanism of action.

[0146] The fourth characteristic involves whether movement of ion species requires energy input. Ion movement down a gradient does not require the input of energy, and is referred to as passive transport or facilitated diffusion. Ion movement up a concentration gradient requires energy, and is referred to as active transport. The energy for active transport may come from a variety of sources. For example, the energy may come from the hydrolysis of ATP. Alternatively, the energy may come from coupling ion movement requiring energy to ion movement that releases energy.

[0147] In addition to the wide variety of channels, pumps, and carriers, gap junctions (GJs) also modulate ion flux across membranes. Gap junctions are plasma membrane protein complexes that directly connect the cytoplasm of neighboring cells, allowing for cell-cell exchange known as gap junctional communication (GJC). GJs can be opened or closed post-translationally (in response to changes in local Ca^{2+} concentration, pH, and/or membrane voltage) and provide for fairly complex gating of the flow of ions and small molecules (<1 kD) directly between cells (Bruzzone and Giaume, 1999, Advances in Experimental Medicine and Biology 468: 321-337; Goldberg et al., 1999, Nature Cell Biology 1: 457-459; Jalife et al., 1999, J of Cardiovascular Electrophysiology 10: 1649-1663; Lampe and Lau, 2000, Archives of Biochemistry and Biophysics 384: 205-215; and White et al., 1994, Nature 371: 208-209). If the GJs connecting two cells are opened, any ion gradient that previously existed will quickly, although not instantaneously, dissipate. Thus, ion flux through GJs can be an important determinant of ion distribution in a tissue, and expression patterns of GJs can establish isopotential cell fields.

[0148] The above characteristics are useful for mechanistically describing how particular ion transporter proteins and classes of ion transporter proteins function to shuttle ion species and regulate ion flux across cell membranes. Many of the identified ion transporter proteins were named based on one or more of their functional characteristics. Throughout this application, however, the generic term ion transporter protein encompasses any of the foregoing categories of proteins (e.g., GJs, channels, pumps, etc) that function to shuttle one or more ion species across cell membranes to mediate ion

flux, membrane potential, and/or pH. Other specific terminology is used, as necessary, to refer to a particular class of ion transporter proteins or to a particular ion transporter protein.

[0149] In addition to the many ion-specific concentration differences (the chemical gradients) that exist as a consequence of membrane flux, there is also a single, all-inclusive charge difference (the electrical gradient) across the membrane. An electrical gradient is called a voltage, and cellular membranes are described as having a membrane voltage (V_m). Many cells, for example differentiated cells, have resting potentials on the order of -60 mV.

[0150] V_m is determined by the gradients, and therefore the concentrations, of all the ions that move across cell membranes. Thus, changes in ion flux via a particular ion transporter protein may alter both the relative concentration of the particular ion in the cell and the membrane potential of the cell. Accordingly, the methods of the present invention can be used to assess not only whether modulation of a particular class of ion transporter proteins modulates ion flux, thereby mediating a particular biological process, but also whether modulation of a particular class of ion transporter proteins modulates membrane voltage, thereby mediating a particular biological process.

[0151] The appreciation that ion flux and membrane potential serve not only a general house keeping function, but also help mediate and/or initiate complex biological events is a recent phenomenon. To illustrate, ion transporter proteins can act to directly regulate the function of non-ion transporter proteins. Alternatively, ion transporter proteins can mediate the flux of an ion species that in turn interacts with an oppositely charged point on a peptide resident within a cell, thereby changing the charge distribution and/or structure of the protein. By way of another example, ion flux can influence downstream events by essentially transforming a cell or tissue into an electrophoresis apparatus (Levin and Mercola, 1998, Developmental Biology 203: 90-105; Levin and Mercola, 1999, Development 126: 4703-4714; Levin et al., 2002, Cell 111: 77-89, Fukumoto, T., Kema, I., and Levin, M., 2005, Current Biology, 15: 794-803; Adams D. S., Robinson K. R., Fukumoto T., Yuan S., Yelick P., Kuo L., McSweeney M., Levin M., 2006, Development, 133: 1657-1671; Hicks, C., Sorocco, D., and Levin, M., 2006, Journal of Neurobiology, in press; Esser, A. T., Smith, K. C., Weaver, J. C., and Levin, M., 2006, Developmental Dynamics, in press; Levin, M., Lauder, J., and Buznikov, G., 2006, Developmental Neuroscience, 28:171-185).

[0152] The invention provides a reiterative method for systematically (i) evaluating whether ion flux is involved in a particular biological process in a particular model system, and, if so, (ii) identifying a broad class of ion transporter proteins that mediate ion flux during the particular biological process and (iii) systematically and progressively narrowing the class of ion transporter proteins that mediate ion flux during the particular biological process. In this way, the method provides a systematic approach for identifying a subset of ion transporter proteins that modulate ion flux during a particular biological process. This subset of ion transporter proteins can then be individually studied to understand the role of individual ion transporter proteins during a particular biological process in a particular model system.

[0153] The logical structure of the screen is as follows. The first compound evaluated for an effect on a particular biological process is a compound with low specificity. In other words, the first compound modulates the activity of a large

family or multiple families of ion transporter proteins (Step 1). We note that at this and any subsequent steps in the assay, compounds are applied in an amount effective to penetrate cells and to modulate ion flux without producing toxic effects. One of skill in the art can readily titrate compounds to select dosages of compounds appropriate for the particular model system.

[0154] If no effect on the particular biological process being assayed is observed after applying the step 1 compound, the entire broad class of ion transporter proteins is eliminated as a candidate for having a role in that particular biological process or model system. We note that, as confirmation of this result, one may optionally repeat the step 1 experiment using a different compound that also modulates the activity of the same broad class of ion transporter proteins. Failure of either compound to produce a detectable change in your biological process provides increased confidence that the broad class of ion transporter proteins is not involved in that biological process or model system.

[0155] If, however, a phenotype (e.g., a change in the particular biological process being observed) is detected following administration of the step 1 compound, the broad class of ion transporter proteins modulated by the step 1 compound is identified as a candidate class of ion transporter proteins that may be involved in mediating the particular biological process. In a second step, a population of cells is contacted with a second compound that is more specific. In other words, the second compound modulates ion flux mediated by a narrower class of ion transporter proteins. This narrower class of ion transporter proteins is a subset of the first, broad class of ion transporter proteins (step 2).

[0156] The method can continue using the same logic. At each subsequent round of screening (step 1, 2, 3, 4, etc), test cells are contacted with an increasingly specific compound. In other words, in each subsequent step, the cells are contacted with a compound that modulates ion flux mediated by an increasingly defined class of ion transporter proteins.

[0157] The number of steps of screening will depend on the system, and on the particular class of ion transporter proteins implicated in the previous step. 2, 3, 4, or more than 4 screening steps may suffice. Once the identified class of ion transporter proteins is sufficiently well defined and of reasonable size, one can begin analyzing individual ion transporter proteins. The number of ion transporter proteins within a class that can be reasonably analyzed individually will vary depending on the resources of the investigator, and the particular model system being used.

[0158] FIG. 1 provides an example of how this methodology can be used to systematically evaluate a role for calcium flux in a particular biological process and/or system. FIG. 1 summarizes a screening assay that includes up to five steps. Note that one can stop screening at any stage and begin analyzing the particular class of ion transporter proteins so identified. Alternatively, additional rounds of screening can be conducted.

[0159] At each step, a compound that modulates ion flux mediated by a class of ion transporter proteins is selected, and cells are contacted with said compound. Exemplary compounds that modulate the activity of a particular class of ion transporter proteins are depicted in the boxes shown to the left of the divider 10. The compounds have increasing specificity for a class of ion transporter proteins as the method progresses through steps 1, 2, 3, 4, and 5. Thus, a step 3 compound modulates a more narrowly defined class of ion transporter proteins than a class 2 compound, and a class 4 compound modulates a more narrowly defined class of ion transporter proteins than a class 3 compound.

proteins then a class 2 compound, and a class 4 compound modulates a more narrowly defined class of ion transporter proteins than a class 3 compound.

[0160] To illustrate, one selects from amongst the compounds depicted in the box under step 1. If one of these compounds produces a phenotype in your particular system, calcium ion flux is identified as involved in your particular system. Thus, in step 1, one has identified an involvement for a broad class of ion transporter proteins that mediate calcium ion flux.

[0161] In step 2, one selects from amongst the compounds depicted in either of the boxes under step 2. If, for example, calciclidine, flunarizine, lamatrigine, lanthanum, riluzole, loperamide hcl, or a functionally similar compound produces a phenotype in your system, the broad class of ion transporter proteins that mediate calcium flux in your system has been narrowed to the calcium channel (CaC) class of transporter proteins. If, on the other hand, gallopamil, NaCN, prenylamine, or thapsigargin produces a phenotype in your system, the broad class of ion transporter proteins that mediate calcium flux in your system has been narrowed to the ATPase class of transporter proteins.

[0162] In each subsequent step, increasingly specific compounds can be used to identify an increasingly more define class of ion transporter proteins that mediate calcium flux, thereby producing a phenotype in your particular biological system.

[0163] We note that, in FIG. 1, the compounds listed to the left of divider 10 are ion transporter inhibitors unless otherwise indicated. Compounds are shown to the left of divider 10 and targets (e.g., class of ion transporter proteins) are represented to the right of divider 10. We also note that the listed compounds are exemplary. Additionally, we note that an investigator may chose to conduct each step of the screen in parallel and test a different compound that functions to inhibit the same class of compounds in parallel populations of cells. Observing the same or a similar phenotype using multiple compounds that modulate the same class of ion transporter proteins can be used to confirm the result and helps control for non-specific effects of the compounds.

[0164] FIG. 2 provides an example of how this methodology can be used to systematically evaluate a role for potassium flux in a particular biological process and/or system. FIG. 2 summarizes a screening assay that may include up to five steps. Note that one can stop screening at any stage and begin analyzing the particular class of ion transporter proteins so identified. Alternatively, additional rounds of screening can be conducted.

[0165] At each step, a compound that modulates ion flux mediated by a class of ion transporter proteins is selected, and cells are contacted with said compound. Exemplary compounds that modulate the activity of a particular class of ion transporter proteins are depicted in the boxes shown to the left of the divider 20. The compounds have increasing specificity for a class of ion transporter proteins as the method progresses through steps 1, 2, 3, 4, and 5. Thus, a step 3 compound modulates a more narrowly defined class of ion transporter proteins than a class 2 compound, and a class 4 compound modulates a more narrowly defined class of ion transporter proteins than a class 3 compound.

[0166] To illustrate, one selects from amongst the compounds depicted in the box under step 1. If one of these compounds produces a phenotype in your particular system, potassium ion flux is identified as involved in your particular

system. Thus, in step 1, one has identified an involvement for a broad class of ion transporter proteins that mediate potassium ion flux.

[0167] In step 2, one selects from amongst the compounds depicted in either of the boxes under step 2. If, for example, Ba, Cs, dimethadione, ergotoxin, tedisamil, YS035, SG209, or a functionally similar compound produces a phenotype in your system, the broad class of ion transporter proteins that mediate potassium flux in your system has been narrowed to the potassium channel class of transporter proteins.

[0168] In each subsequent step, increasingly specific compounds can be used to identify an increasingly more define class of ion transporter proteins that mediate potassium flux, thereby producing a phenotype in your particular biological system.

[0169] We note that, in FIG. 2, the compounds listed to the left of divider 20 are ion transporter inhibitors unless otherwise indicated. Compounds are shown to the left of divider 20 and targets (e.g., class of ion transporter proteins) are represented to the right of divider 20. We also note that the listed compounds are exemplary. Additionally, we note that an investigator may chose to conduct each step of the screen in parallel and test a different compound that functions to inhibit the same class of compounds in parallel populations of cells. Observing the same or a similar phenotype using multiple compounds that modulate the same class of ion transporter proteins can be used to confirm the result and helps control for non-specific effects of the compounds.

[0170] FIG. 3 provides an example of how this methodology can be used to systematically evaluate a role for hydrogen ion flux in a particular biological process and/or system. FIG. 3 summarizes a screening assay that include up to three steps. Note that one can stop screening at any stage and begin analyzing the particular class of ion transporter proteins so identified. Furthermore, one can conduct additional rounds of screening, if desired.

[0171] At each step, a compound that modulates ion flux mediated by a class of ion transporter proteins is selected, and cells are contacted with said compound. At each step, one can select a compound. Exemplary compounds that modulate the activity of a particular class of ion transporter proteins are depicted in the boxes shown to the left of the divider 30. The compounds have increasing specificity for a class of ion transporter proteins as the method progresses through steps 1, 2, 3. Thus, a step 3 compound modulates a more narrowly defined class of ion transporter proteins than a class 2 compound.

[0172] To illustrate, one selects from amongst the compounds depicted in the box under step 1. If one of these compounds produces a phenotype in your particular system, hydrogen ion flux is identified as involved in your particular system. Thus, in step 1, one has identified an involvement for a broad class of ion transporter proteins that mediate hydrogen ion flux.

[0173] In steps 2 and 3, one selects from amongst the compounds to distinct, for example, whether hydrogen ion flux in this particular system is modulated by one of the following classes of ion transporter proteins: V-type H-ATPase, ATP synthase, H/K-ATPase, Na/H exchanger, or H-peptide exchanger.

[0174] In each subsequent step, increasingly specific compounds can be used to identify an increasingly more define class of ion transporter proteins that mediate hydrogen flux, thereby producing a phenotype in your particular biological system.

[0175] We note that, in FIG. 3, the compounds listed to the left of divider 30 are ion transporter inhibitors unless otherwise indicated. Compounds are shown to the left of divider 30 and targets (e.g., class of ion transporter proteins) are represented to the right of divider 30. We also note that the listed compounds are exemplary. Additionally, we note that an investigator may chose to conduct each step of the screen in parallel and test a different compound that functions to inhibit the same class of compounds in parallel populations of cells. Observing the same or a similar phenotype using multiple compounds that modulate the same class of ion transporter proteins can be used to confirm the result and helps control for non-specific effects of the compounds.

[0176] FIG. 4 provides an example of how this methodology can be used to systematically evaluate a role for sodium flux in a particular biological process and/or system. FIG. 4 summarizes a screening assay that may include up to four steps. Note that one can stop screening at any stage and begin analyzing the particular class of ion transporter proteins so identified. Furthermore, one can conduct additional round of screening. At each step, one can select a compound. Exemplary compounds that modulate the activity of a particular class of ion transporter proteins are depicted in the boxes shown to the left of the divider 40. The compounds have increasing specificity for a class of ion transporter proteins as the method progresses through steps 1, 2, 3, and 4. Thus, a step 3 compound modulates a more narrowly defined class of ion transporter proteins then a class 2 compound, and a class 4 compound modulates a more narrowly defined class of ion transporter proteins than a class 3 compound.

[0177] To illustrate, one selects from amongst the compounds depicted in the box under step 1. If one of these compounds produces a phenotype in your particular system, sodium ion flux is identified as involved in your particular system. Thus, in step 1, one has identified an involvement for a broad class of ion transporter proteins that mediate sodium ion flux.

[0178] In step 2, one selects from amongst the compounds depicted in either of the boxes under step 2. If, for example, amiloride hcl, benzamil, lidocaine, or a functionally similar listed or unlisted compound produces a phenotype in your system, the broad class of ion transporter proteins that mediate sodium flux in your system has been narrowed to the sodium channel class of transporter proteins.

[0179] In each subsequent step, increasingly specific compounds can be used to identify an increasingly more define class of ion transporter proteins that mediate sodium flux, thereby producing a phenotype in your particular biological system.

[0180] We note that, in FIG. 4, the compounds listed to the left of divider 40 are ion transporter inhibitors unless otherwise indicated. Compounds are shown to the left of divider 40 and targets (e.g., class of ion transporter proteins) are represented to the right of divider 40. We also note that the listed compounds are exemplary. Additionally, we note that an investigator may chose to conduct each step of the screen in parallel and test a different compound that functions to inhibit the same class of compounds in parallel populations of cells. Observing the same or a similar phenotype using multiple compounds that modulate the same class of ion transporter proteins can be used to confirm the result and helps control for non-specific effects of the compounds.

[0181] FIG. 5 provides an example of how this methodology can be used to systematically evaluate a role for chloride

flux in a particular biological process and/or system. FIG. 5 summarizes a screening assay that may include up to three steps. Note that one can stop screening at any stage and begin analyzing the particular class of ion transporter proteins so identified. Furthermore, one can conduct additional round of screening. At each step, one can select a compound. Exemplary compounds that modulate the activity of a particular class of ion transporter proteins are depicted in the boxes shown to the left of the divider 50. The compounds have increasing specificity for a class of ion transporter proteins as the method progresses through steps 1, 2, and 3.

[0182] To illustrate, one selects from amongst the compounds depicted in the box under step 1. If one of these compounds produces a phenotype in your particular system, chloride ion flux is identified as involved in your particular system. Thus, in step 1, one has identified a role for a broad class of ion transporter proteins that mediate chloride ion flux.

[0183] In step 2, one selects from amongst the compounds depicted in either of the boxes under step 2. If, for example, 9-AC, anthranilic acid, tamoxifen or a functionally similar listed or unlisted compound produces a phenotype in your system, the broad class of ion transporter proteins that mediate chloride flux in your system has been narrowed to the chloride channel class of transporter proteins.

[0184] In each subsequent step, increasingly specific compounds can be used to identify an increasingly more define class of ion transporter proteins that mediate chloride flux, thereby producing a phenotype in your particular biological system.

[0185] We note that, in FIG. 5, the compounds listed to the left of divider 50 are ion transporter inhibitors unless otherwise indicated. Compounds are shown to the left of divider 50 and targets (e.g., class of ion transporter proteins) are represented to the right of divider 50. We also note that the listed compounds are exemplary. Additionally, we note that an investigator may chose to conduct each step of the screen in parallel and test a different compound that functions to inhibit the same class of compounds in parallel populations of cells. Observing the same or a similar phenotype using multiple compounds that modulate the same class of ion transporter proteins can be used to confirm the result and helps control for non-specific effects of the compounds.

[0186] FIG. 6 provides a list of other exemplary compounds that can be used to evaluate whether particular classes of ion transporter proteins are involved in a particular biological process. Any of these drugs can be incorporated into a multi-step screen, or used as part of a single round screen to evaluate the role of a particular class of ion transporter proteins.

[0187] To illustrate, if contacting a population of cells with Cu produces a change in a particular biological process (e.g., a detectable phenotype), aquaporins are identified as having a role in modulating the biological process. If contacting a population of cells with TPEN produces a change in a particular biological process (e.g., a detectable phenotype), Zn, Cu, or Fe ion flux are identified as candidate class of ion transporter proteins.

[0188] FIG. 7 depicts a detailed example of an exemplary multi-step screen for identifying a role for calcium ion flux in a particular biological process, and for further identifying a class of ion transporter proteins that mediate ion flux during the particular biological process. In a first step, a population of cells is contacted with calciclidine. If no effect is observed in the population of cells, the calcium channel class of ion

transporter proteins does not modulate ion flux to mediate the particular biological process in this system. If a phenotype is observed, a second round of screening is conducted.

[0189] In the second round of screening, a second population of equivalent cells is contacted with conotoxin MVIIIC. Depending on whether or not a phenotype is observed, the class of ion transporter protein involved in the biological process, can be further defined in a third, fourth, and fifth round of screening. Ultimately, the results of these rounds of screening can be used to distinguish which of the following class of calcium ion transporter proteins (e.g., T-type, P/Q-type, N-type, R-type, or L-type) mediate ion flux in a particular biological system.

[0190] Table 1 provides a list of exemplary compounds that modulate the activity of one or more classes of ion transporter proteins. Table 1 includes compounds that inhibit the activity of the class of ion transporter proteins, as well as compounds that promote the activity of the class of ion transporter proteins. Compounds that either inhibit or promote the activity of a class of ion transporter proteins can be used to disrupt the endogenous ion flux mediated by a particular class of ion transporter proteins. The existence of numerous compounds that modulate the activity of classes of ion transporters facilitates the screening hierarchy described in the present application. Skilled practitioners can select from amongst available compounds to manipulate ion transporter activity in the biological system being studied.

[0191] The above screening methodology allows an investigator to determine whether ion flux play a role in a particular biological process in a given model system. If so, the method can be used to define a class of ion transporter proteins that mediates ion flux, thereby modulating the biological process. By efficiently focusing the inquiry from amongst the thousands of possible ion transporter proteins to a relatively narrow subset of ion transporter proteins, the screening method allows investigators to define a tractable list of candidate ion transporter proteins for further study.

[0192] Once a class of ion transporter proteins is identified, the resources and techniques of molecular biology can be brought to bear to help identify the particular ion transporter protein involved in the particular biological process and system. For example, the mRNA and protein expression of individual ion transporter proteins within the identified class can be examined (e.g., so called expression analysis). Expression analysis can help determine which members of the identified class of ion transporter proteins are expressed in a spatio-temporal pattern consistent with a role in the particular biological process.

[0193] Alternatively or additionally, functional studies can be conducted in cells or whole organism. Expression or activity of individual ion transporter proteins can be modulated (e.g., inhibited or promoted) to assess whether alteration in the expression or activity of an individual ion transporter protein is sufficient to produce a phenotype. Expression or activity of individual ion transporter proteins can be modulated by administering a compound that specifically alters the expression or activity of an individual ion transporter protein. In certain embodiments, a compound that specifically alters the expression or activity of an ion transporter protein specifically inhibits the expression or activity. In certain other embodiments, a compound that specifically alters the expression or activity of an ion transporter protein specifically promotes the expression or activity. In certain embodiments, a compound that specifically alters the expression or activity of

an ion transporter protein does not substantially alter the expression or activity of other ion transporter proteins in the identified class of ion transporter proteins. In certain other embodiments, a compound that specifically alters the expression or activity of an ion transporter protein does not substantially alter the expression or activity of other classes of ion transporter proteins.

[0194] By way of example, compounds that specifically inhibit or promote the expression or activity of an ion transporter protein include small organic or inorganic molecules. By way of further example, compounds that specifically inhibit the expression or activity of an ion transporter protein include nucleic acids (e.g., sense or antisense oligonucleotide, RNAi constructs, etc). By way of further example, compounds that specifically inhibit the expression or activity of an ion transporter protein include proteins (e.g., antibodies, polypeptides). By way of further example, compounds that specifically promote the expression or activity of an ion transporter protein include nucleic acids or proteins. By way of further example, compounds that specifically promote the expression or activity of an ion transporter protein include nucleic acid expression constructs of the ion transporter. Further embodiments are described in section ii.

[0195] Functional studies of individual ion transporter proteins can also be facilitated by expressing candidate ion transporter proteins as fusions proteins with a fluorescent protein such as GFP. This allows observation of the trafficking and subcellular localization of individual ion transporter proteins.

[0196] Additionally, individual ion transporter proteins or classes of ion transporter proteins can also be assessed to determine whether they modulate membrane potential and/or pH in the particular biological process. Furthermore, ion flux can be directly examined. The invention contemplates that direct detection of ion flux, intracellular pH, and/or membrane potential can be used at any stage of the screening and characterization process. In other words, such direct measurements of biophysical phenomenon can be conducted early in the screening process during evaluation of larger candidate classes of ion transporter proteins. Alternatively, direct measurements can be conducted when characterizing individual ion transporter proteins. In other embodiments, direct measurements of biophysical characteristics are not performed. Further embodiments are described in section v.

[0197] The foregoing methods can be conducted in cells in culture, in tissue samples maintained ex vivo, or in animals. When the method is conducted using cells in culture, the invention contemplates using cells derived from any organism, tissue, or stage of development. Furthermore, the invention contemplates that the cells may be primary cultures of cells, or transformed cell lines, and that the cells can either be wild type cells or cells containing one or more mutations. Mutant cells or cell lines may be models of a particular disease or injury, or may be derived from animals having a specific disease or injury (e.g. cancer cells harvested from an animal).

[0198] Cells may be derived from (e.g., derived from and cultured in vitro as populations of cells or tissues) or reside in (cells resident in a whole animal or portion of a whole animal) any of a number of animal species. Exemplary animals include, but are not limited to, flatworms, amphibians, fish, reptiles, birds, or mammals. Suitable flatworms include planarian. Suitable amphibians include *Xenopus laevis*, *Xenopus tropicalis*, and other species of frog. Suitable birds include chickens, as well as other birds commonly used or maintained

in a laboratory setting. Suitable mammals include mice, rats, hamsters, goats, sheep, pigs, cows, dogs, cats, rabbits, non-human primates, and humans.

[0199] Regardless of the species of cells or animal selected, the invention contemplates that cells may be derived from or reside in an animal of virtually any stage of development. For example, the cells may be derived from or reside in an embryonic, larval, fetal, juvenile, or adult organism. The decision of whether to conduct a particular screen in cells, tissues, or animals, as well as the species and stage of development of the selected cells or animals can be readily made by one of skill in the art. The skilled artisan can select the approach, conditions, and system based on their expertise, resources, and the particular biological process they are investigating. Further embodiments are described in section iii.

[0200] In addition, the foregoing methods can be used to evaluate a role for ion flux mediated by a class of ion transporter proteins in virtually any biological process or system. By way of example, the methods can be used to evaluate a role for ion flux in cell proliferation, cell differentiation, cell survival, cell apoptosis, cell dedifferentiation, cell regeneration, or cell migration. These and other biological processes can be observed/assayed using morphological criteria (e.g., visual inspection) or the biological processes can be observed using cell biological or molecular tools and reagents. For example, the biological process can be assessed by detecting a change in gene expression (e.g., transcription or translation) using one or more molecular markers, a change in rate of cell proliferation (e.g., BrdU incorporation), a change in apoptosis (TUNEL analysis), and the like.

[0201] The foregoing represents a powerful strategy for using known pharmacological compounds to rapidly and inexpensively implicate and study specific candidates proteins for roles in any biological process (Adams and Levin, Chapter 9, Sater and Whitman, 2006). Although individual pharmacological drugs have been used previously in studies of toxicology, neuropharmacology, and teratology, these studies have failed to provide a hierarchical process whereby a role for particular proteins and signaling pathways can be efficiently evaluated.

[0202] Previous drug-based screens have generally fallen into two categories: exhaustive screening of large numbers of drugs (so-called "Sigma screens") or specific testing of single compounds or small numbers of compounds to confirm the role of a target that has already been identified as a candidate via some other means. These approaches are fundamentally different from the screening system provided in the present application. Furthermore, approaches previously used in the art have serious limitations which are overcome by the system proposed in the present application.

[0203] The term "Sigma screens" is often used to refer to a large scale, exhaustive screening of compounds. The term is derived from the practice of screening through compounds available commercially through a major supplier such as Sigma. Usually, large numbers of compounds are tested. Although there may be an assay in which the function or efficacy of the compounds are evaluated, these screens did not typically provide any mechanism for relating effects (positive or negative) obtained using one compound with those obtained using another compound. In other words, these screens have not been performed in a hierarchical fashion which either relates effects across compounds or helps the scientist make decisions about what other compounds should be tested. In sharp contrast, the present invention provides a

hierarchical approach that both relates information and helps guide the practitioner in selecting other compounds for testing.

[0204] Candidate screening approaches can be very useful to study a protein, pathway, or process that is already understood or for which considerable data already exists. However, this approach is laborious and time consuming for studying processes, proteins, or pathways for which considerable information is not already available. Furthermore, candidate screening approaches do not provide any way to interrelate the data obtained when testing individual candidate.

[0205] The approach of this application systematizes the screening process and provides a hierarchical approach to uncovering the role of molecular players (proteins and pathways) in interesting developmental events and biological processes. This approach described above, provides a highly efficient way to systematically identify and study the roles for proteins and pathways in a wide range of biological processes and systems.

[0206] (c) Methods of Identifying and/or Purifying Progenitor Cells

[0207] The present invention further provides methods for identifying and/or purifying progenitor cells from amongst heterogeneous populations of cells or from organisms. This aspect of the invention is based on the recognized correlation between the level of cell differentiation (e.g., how committed along a particular cell lineage is a particular cell?) and the membrane potential of that cell. This correlation can be used to identify cells among a population of cells whose membrane voltage characteristics are most consistent with a progenitor cell-like state. Identified cells can then be further cultured and/or studied to (i) confirm that the identified cells possess other characteristics of progenitor cells, (ii) purify the progenitor cells, and/or (iii) expand the progenitor cells.

[0208] Progenitor cells exist during embryonic, fetal, and adult development. There study has generated enormous scientific interest because progenitor cells are believed to provide promising potential treatments for a range of degenerative diseases and injuries. Furthermore, the study of progenitor cells will likely increase our understanding of normal development and regenerative processes.

[0209] The term "progenitor cell" is used synonymously with "stem cell". Both terms refer to an undifferentiated cell which is capable of proliferation and giving rise to more progenitor cells having the ability to generate a large number of mother cells that can in turn give rise to differentiated, or differentiable daughter cells. In a preferred embodiment, the term progenitor or stem cell refers to a generalized mother cell whose descendants (progeny) specialize, often in different directions, by differentiation, e.g., by acquiring completely individual characters, as occurs in progressive diversification of embryonic cells and tissues.

[0210] The term "embryonic stem cell" is used to refer to the pluripotent stem cells of the inner cell mass of the embryonic blastocyst (see U.S. Pat. Nos. 5,843,780, 6,200,806). Such cells can similarly be obtained from the inner cell mass of blastocysts derived from somatic cell nuclear transfer (see, for example, U.S. Pat. Nos. 5,945,577, 5,994,619, 6,235, 970).

[0211] The term "adult stem cell" is used to refer to any multipotent stem cell derived from tissues other than the embryonic blastocyst. Adult stem cells include cells derived from non-blastocyst tissue, including tadpole, fetal, juvenile, and adult tissue. Stem cells have been isolated from a wide

variety of adult tissues (e.g., non-blastocyst) including blood, bone marrow, brain, olfactory epithelium, skin, pancreas, skeletal muscle, and cardiac muscle. Each of these stem cells can be characterized based on gene expression, factor responsiveness, and morphology in culture. Exemplary adult stem cells include neural stem cells, neural crest stem cells, mesenchymal stem cells, hematopoietic stem cells, and pancreatic stem cells. As indicated above, stem cells have been found resident in virtually every tissue. Accordingly, the invention contemplates the identification of progenitor cells resident in any tissue, in any organism, during any stage of development.

[0212] The term "substantially pure", with respect to a particular cell population, refers to a population of cells that is at least about 75%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% pure, with respect to the cells making up a total cell population. Recast, the term "substantially pure" refers to a population of cells that contain fewer than about 20%, more preferably fewer than about 10%, most preferably fewer than about 5%, of lineage committed cells.

[0213] Research using progenitor cells has been hampered by a number of issues. Amongst the technical issues are the difficulty in readily identifying and isolating progenitor cells from amongst heterogeneous cell populations *in vivo* or *in vitro*. Progenitor cells are generally believed to constitute a very small percentage of the total number of cells in a given tissue, and even many available culture techniques still produce heterogeneous populations of cells containing both progenitor and non-progenitor cells. Difficulty in identify progenitor cells and difficulty in obtaining purified cultures of progenitor cell populations hampers further study, and may ultimately hamper the development of effective therapies based on progenitor cells.

[0214] The present invention provides methods of identifying progenitor cells from amongst heterogeneous populations of cells. Identified progenitor cells can then be separated or otherwise removed from amongst the heterogeneous population of cells. The separated cells can then be cultured to generated substantially purified populations of progenitor cells.

[0215] The methods of the invention are amenable to use with cells *in vitro* (e.g., to identify progenitor cells from amongst heterogeneous populations of cells in culture). The methods of the invention are amenable to use with cells *in vivo* (e.g., to identify progenitor cells resident in the tissues of animals or animal fragments). The methods of the invention can be used to identify progenitor cells present during any stage of development in any tissue type. As such, the present methods provide an approach of broad applicability to research across the stem cells field.

[0216] The identification and purification methods of the invention are based on the electrogenic properties of cells. Terminally differentiated cells and other non-proliferating cells generally have membrane potentials of less than -50 mV. For example, differentiated neurons have a membrane potential of approximately -90 mV, differentiated skeletal muscle has a membrane potential of approximately -75 mV, fat cells have a membrane potential of approximately -70 mV, kidney tubules have a membrane potential of approximately -65 mV, and smooth muscle cells have a membrane potential of approximately -60 mV. In contrast, proliferating cells including tumor cells and blastomeres of early cleavage stage embryos are depolarized. By way of example, blas-

tomeres of a 16 cell embryo have a membrane potential of approximately -25 mV and the fertilized egg has a membrane potential of approximately -10 mV.

[0217] In addition to observations regarding the relative membrane potential of differentiated cells versus proliferating uncommitted cells, a correlation between intracellular pH and cell commitment also exists. Differentiated cells appear to have a higher intracellular pH in comparison to proliferating cells such as cancer cells. For example, a survey of several differentiated cell types recorded intracellular pH levels of between about 6.9-7.4. In contrast, intracellular pH levels of less than or equal to about 6.8 are consistently observed in various tumor cells in culture.

[0218] Based on the above correlations, as well as our observation of depolarized membrane potential during regeneration in model systems including planaria and frog, the present invention provides a generalized method for identifying progenitor cells from amongst heterogeneous populations of cells. The present invention provides methods for identifying progenitor cells. The method for identifying progenitor cells comprises contacting a population of cells with a voltage sensitive agent. The voltage sensitive agent produces a detectable signal in the presence of cells having a depolarized cell membrane. In this way, the voltage sensitive agent identifies the cells, if any, with a depolarized cell membrane. Cells having a depolarized cell membrane are identified, and these cells can be separated from the heterogeneous population of cells. The separated cells can be further studied and cultured. For example, separated cells can be analyzed with molecular markers indicative of particular progenitor cell populations to confirm that the cells are progenitor cells.

[0219] Based on previously observed correlations between membrane potential, one of skill in the art can select the appropriate membrane potential for the selection and separation of cells for further analysis. In one embodiment, cells having a depolarized cell membrane with a membrane potential of greater than or equal to -20 mV are identified. Such cells are candidate progenitor cells that are separated from the heterogeneous population of cells and further analyzed and cultured. In other embodiments, cells having a depolarized cell membrane with a membrane potential of greater than or equal to -15 mV, -10 mV, -5 mV, -3 mV, 0 mV, 5 mV, 10 mV, 15 mV, or mV are identified. Such cells are candidate progenitor cells that are separated from the heterogeneous population of cells and further analyzed and cultured.

[0220] The invention contemplates that the optimal membrane potential cut-off for effectively identifying progenitor cells without also including substantial numbers of non-progenitor cells may vary across tissues and organisms. For example, the optimal condition for identifying substantial numbers of progenitor cells while including a limited percentage of non-progenitor cells may differ across tissues and organisms. However, one of skill in the art can readily select the appropriate cut-off. The important point is that the first step of identifying progenitor cells allow one to separate a group of cells from a heterogeneous population of cells and substantially enrich for progenitor cells. Although the first step can produce a substantially purified population of progenitor cells, it need not do so. Even when the first step provides an initial identification of cells that include both progenitor cells and some non-progenitor cells, the method is still useful for producing populations of cells that are substantially enriched for progenitor cells. Additional rounds of screening and/or examination of known progenitor cell mark-

ers can be used, as necessary, to further enrich for progenitor cells or to confirm that a population of cells comprises progenitor cells.

[0221] As outlined above, intracellular pH is also correlated with the differentiation state and regenerative capacity of a cell. Accordingly, the foregoing methods can use, alternatively or in addition to, detection of intracellular pH to identify progenitor cells amongst a heterogeneous population of cells. In one embodiment, cells having an intracellular pH of cells than or equal to about 6.7 are identified for further study (e.g., separation, culture, further purification). In another embodiment, cells having an intracellular pH of less than or equal to about 6.6, 6.5, 6.4, 6.3, 6.2, 6.1, or 6.0 are identified for further study.

[0222] Detection of membrane voltage or intracellular pH can be used alone or in combination in the subject methods of identifying progenitor cells.

[0223] Methods and agents for detecting membrane voltage and pH are described in detail above. Numerous agents and methods exist and can be used in methods for identifying progenitor cells. By way of non-limited example, voltage sensitive agents include fluorescent dyes. Exemplary dyes include, but are not limited to, bis-(1,3-dibutylbarbituric acid) pentamethine oxonol (DiBAC₄(5)); bis-(1,3-dibutylbarbituric acid)trimethine oxonol (DiBAC₄(3)); bis-(1,3-diethylthiobarbituric acid)trimethine oxonol (DiSBAC₂(3)); 3,3'-diethyloxacarbocyanine iodide (DiOC₂(3)); 3,3'-diheptyloxacarbocyanine iodide (DiOC₇(3)); 3,3'-dihexyloxacarbocyanine iodide (DiOC₆(3)); 3,3'-dipentyloxacarbocyanine iodide (DiOC₅(3)); 1,1',3,3,3',3'-hexamethylindodicarbocyanine iodide (DiIC₁(5)); a structural variant thereof; or a functional variant thereof. Exemplary pH sensitive agents include fluorescent dyes. Such dyes include, but are not limited to, 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein (BCECF); 5-(and-6)-carboxy SNARF®-1; Lysotracker® Blue DND-22; Lysotracker® Green DND-26; Lysotracker® Red DND-99; Lysotracker® Yellow HCK-123; a structural variant thereof; or a functional variant thereof.

[0224] These and other methods and agents for detecting membrane voltage and pH are available and can be adapted for use in identifying progenitor cells in cells, tissues, or organisms. The appropriate methods and agents can be selected based on the organism and stage of development being examined. In addition to fluorescent agents such as pH or voltage sensitive dyes, the invention contemplates that other methods for identifying cells having a particular membrane potential or pH can be used to identify candidate progenitor cells. For example, patch-clamp or vibrating probes methods can be used.

[0225] In certain embodiments, cells identified as having a particular membrane potential and/or pH are separated for further culture and analysis. Separating the identified cells can involve dissecting the cells away from the microarchitecture in which they reside. For example, when the method is conducted in an organism, identified cells can be dissected out of the three-dimensional structures of the whole organism or tissue. Specific methods of microdissection can be selected based on the particular organism used, as well as the number of cells to be removed. Laser-based methods, as well as manual methods employing scalpel, tungsten needles, and other dissection tools can be readily employed.

[0226] When the methods of identification are conducted using cells cultured in vitro, separating identified cells may

employ any of a number of methods. Identified cells can be dissected from the culture. Alternatively, the culture of cells can be dissociated, and the dissociated cells can be separated based on, for example, the detectable agent used to mark one or more cells as candidate progenitor cells. In certain embodiments, the culture of cells is dissociated, and automated sorting methods are used to separate the cells based on the detectable agent. Exemplary automated methods include FACS scan analysis.

[0227] Once progenitor cells are identified, they can be cultured under conditions appropriate for maintaining progenitor cells derived from the particular organism and tissue. For example, media and culture conditions for maintaining mammalian mesenchymal stems cells in an undifferentiated state are known. Similarly, appropriate conditions and reagents for maintaining proliferating cultures of various progenitor cell populations are known. Such methods can be readily employed and adapted during the further study and analysis of progenitor cells identified by the present methods.

(ii) Compounds

[0228] The present invention provides screening methods for evaluating the role of ion flux during a biological process. The present invention further provides methods for identifying and/or purifying progenitor cells. The present invention also provides methods for promoting or inhibiting dedifferentiation and/or regeneration in cells derived from or resident in an organism, as well as methods for screening to identify compounds that promote or inhibit dedifferentiation and/or regeneration, and pharmaceutical preparations comprising the identified compounds. Many of the methods and compositions of the present invention involve contacting populations of cells with compounds.

[0229] The invention contemplates the use of any of a wide range of compounds in the methods and screening assays of the invention. Exemplary classes of compounds include, but are not limited to, nucleic acids, peptides, polypeptides, small organic molecules, small inorganic molecules, peptidomimetics, antisense oligonucleotides, RNAi constructs, ribozymes, and antibodies. Compounds can be screened as single agents, multiple candidate agents, or libraries of agents. Exemplary classes of compounds are described in detail below. Tables 1 and 2 provide numerous available compounds that modulate the activity of particular ion transporters or classes of ion transporters.

[0230] Exemplary compounds include compounds that modulate ion flux by inhibiting the activity of an ion transporter protein or class of ion transporter proteins. Exemplary compounds also include compounds that modulate ion flux by promoting the activity of an ion transporter protein or class of ion transporter proteins. In certain embodiments, compounds that modulate ion flux by either promoting or inhibiting the activity of an ion transporter protein or class of ion transporter proteins also modulate membrane potential. In certain embodiments, compounds that modulate ion flux by either promoting or inhibiting the activity of an ion transporter protein or class of ion transporter proteins can be used to modulate dedifferentiation and/or regeneration.

[0231] Any of the classes of compounds can be formulated and administered as a composition or as a pharmaceutical composition.

[0232] Polypeptides and peptide fragments: In certain embodiments, the compounds are polypeptides or peptide fragments. Exemplary polypeptides or peptide fragments

include wildtype, as well as variant sequences. Variant polypeptides include amino acid sequences at least 60%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% identical to a particular wild type polypeptide.

[0233] In addition to polypeptides and peptide fragments, the present invention also contemplates isolated nucleic acids comprising nucleotide sequences that encode said polypeptides and fragments. The term nucleic acid as used herein is intended to include fragments as equivalents, wherein such fragments have substantially the same function as the full length nucleic acid sequence from which it is derived. Equivalent nucleotide sequences will include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include sequences that differ from the nucleotide sequence of, for example, the native nucleotide sequence. Equivalent sequences include those that vary from a known wildtype or variant sequence due to the degeneracy of the genetic code. Equivalent sequences may also include nucleotide sequences that hybridize under stringent conditions (i.e., equivalent to about 20-27° C. below the melting temperature (T_m) of the DNA duplex formed in about 1M salt) to the native nucleotide sequence. Further examples of stringent hybridization conditions include a wash step of 0.2xSSC at 65° C. Equivalent nucleotide sequences will be understood to encode polypeptides which retain the activity of the polypeptide encoded by the native nucleotide sequence.

[0234] Equivalent nucleotide sequences for use in the methods described herein also include sequences which are at least 60% identical to a given nucleotide sequence. In another embodiment, the nucleotide sequence is at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% identical to the nucleotide sequence of a native sequence.

[0235] Nucleic acids having a sequence that differs from nucleotide sequences which encode a particular polypeptide due to degeneracy in the genetic code are also within the scope of the invention. Such nucleic acids encode functionally equivalent peptides but differ in sequence from wildtype sequences known in the art due to degeneracy in the genetic code. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC each encode histidine) may result in "silent" mutations which do not affect the amino acid sequence. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences will also exist. One skilled in the art will appreciate that these variations in one or more nucleotides (up to about 3-5% of the nucleotides) of the nucleic acids encoding polypeptides may exist among individuals of a given species due to natural allelic variation.

[0236] In certain embodiments, the compound is a nucleic acid sequence that encodes an ion transporter protein. In certain other embodiments, the compound is a polypeptide corresponding to an ion transporter protein. In certain other embodiments, the compound is a nucleic acid sequence that encodes a portion of an ion transporter protein, and the portion of the ion transporter protein acts as a dominant negative construct that inhibits the expression or activity of an endogenous protein. In certain other embodiments, the compound is a polypeptide corresponding to a portion of an ion transporter protein, and the portion of the ion transporter protein acts as a dominant negative construct that inhibits the expression or activity of an endogenous protein. In certain embodiments, the compound is a nucleic acid sequence that encodes a gain

or loss of function mutant ion transporter protein. In certain other embodiments, the compound is a polypeptide corresponding to a gain or loss of function mutant ion transporter protein. In certain embodiments, the compound is a nucleic acid sequence that encodes the V-ATPase H⁺ ion transporter protein. In certain other embodiments, the compound is a polypeptide corresponding to the V-ATPase H⁺ ion transporter protein. In certain embodiments, the compound is a nucleic acid sequence that encodes the P-type H⁺ ATPase. In certain other embodiments, the compound is a polypeptide corresponding to the P-type H⁺ ATPase. Ion transporter proteins or nucleic acid sequences encoding them may be of any species or genotype.

[0237] Antibodies: Exemplary compounds also include antibodies. Antibodies can have extraordinary affinity and specificity for particular epitopes. Without being bound by theory, antibodies can inhibit or potentiate the activity of proteins and signaling pathways in cells, thereby exerting or inducing a particular effect on cells, tissues, or organisms.

[0238] Monoclonal or polyclonal antibodies can be made using standard protocols (See, for example, Antibodies: A laboratory manual ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of a peptide. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of a protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. We note that antibodies may be immunospecific for a particular protein, may be immunospecific for a particular family of proteins, or may be less immunospecific and cross-react with multiple protein from related families of proteins. Antibodies which are immunospecific do not substantially cross-react with non-homologous protein. By not substantially cross react is meant that the antibody has a binding affinity for a non-homologous proteins which is at least one order of magnitude, more preferably at least 2 orders of magnitude, and even more preferably at least 3 orders of magnitude less than the binding affinity of the antibody for the protein or proteins for which the antibody is immunospecific.

[0239] The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with a polypeptide or family of polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab)2 fragments can be generated by treating antibody with pepsin. The resulting F(ab)2 fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibodies of the present invention are further intended to include bispecific and chimeric molecules having affinity for a protein conferred by at least one CDR region of the antibody.

[0240] In one variation, antibodies of the invention can be single chain antibodies (scFv), comprising variable antigen binding domains linked by a polypeptide linker. Single chain antibodies are expressed as a single polypeptide chain and can be expressed in bacteria and as part of a phage display library. The nucleic acid encoding the single chain antibody can then be recovered from the phage and used to produce large quantities of the scFv. Construction and screening of scFv libraries

is extensively described in various publications (U.S. Pat. Nos. 5,258,498; 5,482,858; 5,091,513; 4,946,778; 5,969,108; 5,871,907; 5,223,409; 5,225,539).

[0241] The technology for producing monoclonal antibodies is well known. The preferred antibody homologs contemplated herein can be expressed from intact or truncated genomic or cDNA or from synthetic DNAs in prokaryotic or eukaryotic host cells. The dimeric proteins can be isolated from the culture media and/or refolded and dimerized in vitro to form biologically active compositions. Heterodimers can be formed in vitro by combining separate, distinct polypeptide chains. Alternatively, heterodimers can be formed in a single cell by coexpressing nucleic acids encoding separate, distinct polypeptide chains. See, for example, WO93/09229, or U.S. Pat. No. 5,411,941, for several exemplary recombinant heterodimer protein production protocols. Currently preferred host cells include, without limitation, prokaryotes including *E. coli*, or eukaryotes including yeast, *Saccharomyces*, insect cells, or mammalian cells, such as CHO, COS or BSC cells. One of ordinary skill in the art will appreciate that other host cells can be used to advantage.

[0242] Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using 1500 molecular weight polyethylene glycol ("PEG 1500"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridomas producing a desired antibody are detected by screening the hybridoma culture supernatants.

[0243] To produce antibody homologs that are intact immunoglobulins, hybridoma cells that tested positive in such screening assays were cultured in a nutrient medium under conditions and for a time sufficient to allow the hybridoma cells to secrete the monoclonal antibodies into the culture medium. Tissue culture techniques and culture media suitable for hybridoma cells are well known.

[0244] Alternatively, the desired antibody may be produced by injecting the hybridoma cells into the peritoneal cavity of an unimmunized mouse. The hybridoma cells proliferate in the peritoneal cavity, secreting the antibody which accumulates as ascites fluid. The antibody may be harvested by withdrawing the ascites fluid from the peritoneal cavity with a syringe.

[0245] Fully human monoclonal antibody homologs are another compound that can be used. In their intact form these may be prepared using in vitro-primed human splenocytes, as described by Boerner et al., 1991, J. Immunol., 147, 86-95. Alternatively, they may be prepared by repertoire cloning as described by Persson et al., 1991, Proc. Nat. Acad. Sci. USA, 88: 2432-2436 or by Huang and Stollar, 1991, J. Immunol. Methods 141, 227-236. U.S. Pat. No. 5,798,230 describes preparation of human monoclonal antibodies from human B cells.

[0246] In yet another method for producing fully human antibodies, U.S. Pat. No. 5,789,650 describes transgenic non-human animals capable of producing heterologous antibodies and transgenic non-human animals having inactivated endogenous immunoglobulin genes.

[0247] Large nonimmunized human phage display libraries may also be used to isolate high affinity antibodies that can be developed as human therapeutics using standard phage technology (Vaughan et al., 1996).

[0248] Yet another preferred binding agent is a humanized recombinant antibody homolog. Following the early methods for the preparation of true "chimeric antibodies" (where the entire constant and entire variable regions are derived from different sources), a new approach was described in EP 0239400 (Winter et al.) whereby antibodies are altered by substitution (within a given variable region) of their complementarity determining regions (CDRs) for one species with those from another. The process for humanizing monoclonal antibodies via CDR "grafting" has been termed "reshaping". (Riechmann et al., 1988, Nature 332, 323-327; Verhoeven et al., 1988, Science 239, 1534-1536).

[0249] Antisense, ribozyme and triplex techniques: Nucleic acid-based compounds include, but are not limited to, antisense oligonucleotides and ribozymes. Antisense oligonucleotides and ribozymes inhibit the expression of a protein, e.g., by inhibiting transcription and/or translation.

[0250] Binding of the oligonucleotide or ribozyme to the nucleic acid encoding the particular protein may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, "antisense" therapy refers to the range of techniques generally employed in the art, and includes any therapy that relies on specific binding to oligonucleotide sequences.

[0251] An antisense construct of the present invention can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a particular protein. Alternatively, the antisense construct is an oligonucleotide probe that is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences encoding a particular protein. Such oligonucleotide probes are preferably modified oligonucleotides that are resistant to endogenous nucleases, e.g., exonucleases and/or endonucleases, and are therefore stable in vivo. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Pat. Nos. 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by Van der Krol et al. (1988) BioTechniques 6:958-976; and Stein et al. (1988) Cancer Res 48:2659-2668. With respect to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the nucleotide sequence of interest, are preferred.

[0252] Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to mRNA encoding a particular protein. The antisense oligonucleotides will bind to the mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. In the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex,

as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

[0253] Oligonucleotides that are complementary to the 5' end of the mRNA, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have recently been shown to be effective at inhibiting translation of mRNAs as well. (Wagner, R. 1994. Nature 372:333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of a gene could be used in an antisense approach to inhibit translation of that mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could also be used in accordance with the invention. Whether designed to hybridize to the 5', 3' or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably less than about 100 and more preferably less than about 50, 25, 17 or 10 nucleotides in length.

[0254] The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. WO88/09810, published Dec. 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. WO89/10134, published Apr. 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

[0255] The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxyethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxymaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

[0256] The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

[0257] The antisense oligonucleotide can also contain a neutral peptide-like backbone. Such molecules are termed peptide nucleic acid (PNA)-oligomers and are described, e.g., in Perry-O'Keefe et al. (1996) Proc. Natl. Acad. Sci. U.S.A. 93:14670 and in Egloff et al. (1993) Nature 365:566. One advantage of PNA oligomers is their capability to bind to complementary DNA essentially independently from the ionic strength of the medium due to the neutral backbone of the DNA. In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphordithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

[0258] In yet a further embodiment, the antisense oligonucleotide is an -anomeric oligonucleotide. An -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual-units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

[0259] Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451).

[0260] While antisense nucleotides complementary to the coding region of an mRNA sequence can be used, those complementary to the transcribed untranslated region and to the region comprising the initiating methionine are most preferred.

[0261] The antisense molecules can be delivered to cells or animals *in vitro* or *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systematically. We note that these and other methods are have been used to deliver single antisense oligonucleotides, as well as libraries of oligonucleotides.

[0262] However, it may be difficult to achieve intracellular concentrations of the antisense sufficient to suppress translation on endogenous mRNAs in certain instances. Therefore a preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfet target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous transcripts and thereby prevent translation. For example, a vector can be introduced *in vivo* such that it is

taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Benoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al, 1982, Nature 296:39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct that can be introduced directly into the tissue site. Alternatively, viral vectors can be used which selectively infect the desired tissue, in which case administration may be accomplished by another route (e.g., systematically).

[0263] Ribozyme molecules designed to catalytically cleave an mRNA transcript can also be used to prevent translation of mRNA (See, e.g., PCT International Publication WO90/11364, published Oct. 4, 1990; Sarver et al., 1990, Science 247:1222-1225 and U.S. Pat. No. 5,093,246). While ribozymes that cleave mRNA at site-specific recognition sequences can be used to destroy particular mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988, Nature, 334:585-591.

[0264] The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, 224:574-578; Zaug and Cech, 1986, Science, 231:470-475; Zaug, et al., 1986, Nature, 324:429-433; published International patent application No. WO88/04300 by University Patents Inc.; Been and Cech, 1986, Cell, 47:207-216). The Cech-type ribozymes have an eight base pair active site that hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes that target eight base-pair active site sequences.

[0265] As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and can be delivered *in vivo* or *in vitro*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy targeted messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

[0266] Alternatively, endogenous gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene (i.e., the promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells in the body. (See generally, Helene, C. 1991, *Anticancer Drug Des.*, 6(6): 569-84; Helene, C., et al., 1992, *Ann. N.Y. Acad. Sci.*, 660: 27-36; and Maher, L. J., 1992, *Bioassays* 14(12):807-15).

[0267] Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides should promote triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which will result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules may be chosen that are purine-rich, for example, containing a stretch of G residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

[0268] Antisense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0269] Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

[0270] RNAi: In other embodiments, the compound is an RNAi construct. RNAi constructs comprise double stranded RNA that can specifically block expression of a target gene. "RNA interference" or "RNAi" is a term initially applied to a phenomenon observed in plants and worms where double-stranded RNA (dsRNA) blocks gene expression in a specific and post-transcriptional manner. Without being bound by theory, RNAi appears to involve mRNA degradation, however the biochemical mechanisms are currently an active area of research. Despite some mystery regarding the mechanism of action, RNAi provides a useful method of inhibiting gene expression *in vitro* or *in vivo*.

[0271] As used herein, the term "dsRNA" refers to siRNA molecules, or other RNA molecules including a double

stranded feature and able to be processed to siRNA in cells, such as hairpin RNA moieties.

[0272] The term "loss-of-function," as it refers to genes inhibited by the subject RNAi method, refers to a diminishment in the level of expression of a gene when compared to the level in the absence of RNAi constructs.

[0273] As used herein, the phrase "mediates RNAi" refers to (indicates) the ability to distinguish which RNAs are to be degraded by the RNAi process, e.g., degradation occurs in a sequence-specific manner rather than by a sequence-independent dsRNA response, e.g., a PKR response.

[0274] As used herein, the term "RNAi construct" is a generic term used throughout the specification to include small interfering RNAs (siRNAs), hairpin RNAs, and other RNA species which can be cleaved *in vivo* to form siRNAs. RNAi constructs herein also include expression vectors (also referred to as RNAi expression vectors) capable of giving rise to transcripts which form dsRNAs or hairpin RNAs in cells, and/or transcripts which can produce siRNAs *in vivo*.

[0275] "RNAi expression vector" (also referred to herein as a "dsRNA-encoding plasmid") refers to replicable nucleic acid constructs used to express (transcribe) RNA which produces siRNA moieties in the cell in which the construct is expressed. Such vectors include a transcriptional unit comprising an assembly of (1) genetic element(s) having a regulatory role in gene expression, for example, promoters, operators, or enhancers, operatively linked to (2) a "coding" sequence which is transcribed to produce a double-stranded RNA (two RNA moieties that anneal in the cell to form an siRNA, or a single hairpin RNA which can be processed to an siRNA), and (3) appropriate transcription initiation and termination sequences. The choice of promoter and other regulatory elements generally varies according to the intended host cell. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

[0276] The RNAi constructs contain a nucleotide sequence that hybridizes under physiologic conditions of the cell to the nucleotide sequence of at least a portion of the mRNA transcript for the gene to be inhibited (i.e., the "target" gene). The double-stranded RNA need only be sufficiently similar to natural RNA that it has the ability to mediate RNAi. Thus, the invention has the advantage of being able to tolerate sequence variations that might be expected due to genetic mutation, strain polymorphism or evolutionary divergence. The number of tolerated nucleotide mismatches between the target sequence and the RNAi construct sequence is no more than 1 in 5 basepairs, or 1 in 10 basepairs, or 1 in 20 basepairs, or 1 in 50 basepairs. Mismatches in the center of the siRNA duplex are most critical and may essentially abolish cleavage of the target RNA. In contrast, nucleotides at the 3' end of the siRNA strand that is complementary to the target RNA do not significantly contribute to specificity of the target recognition.

[0277] Sequence identity may be optimized by sequence comparison and alignment algorithms known in the art (see Gribskov and Devereux, Sequence Analysis Primer, Stockton

Press, 1991, and references cited therein) and calculating the percent difference between the nucleotide sequences by, for example, the Smith-Waterman algorithm as implemented in the BESTFIT software program using default parameters (e.g., University of Wisconsin Genetic Computing Group). Greater than 90% sequence identity, or even 100% sequence identity, between the inhibitory RNA and the portion of the target gene is preferred. Alternatively, the duplex region of the RNA may be defined functionally as a nucleotide sequence that is capable of hybridizing with a portion of the target gene transcript (e.g., 400 mM NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA, 50° C. or 70° C. hybridization for 12-16 hours; followed by washing).

[0278] Production of RNAi constructs can be carried out by chemical synthetic methods or by recombinant nucleic acid techniques. Endogenous RNA polymerase of the treated cell may mediate transcription in vivo, or cloned RNA polymerase can be used for transcription in vitro. The RNAi constructs may include modifications to either the phosphate-sugar backbone or the nucleoside, e.g., to reduce susceptibility to cellular nucleases, improve bioavailability, improve formulation characteristics, and/or change other pharmacokinetic properties. For example, the phosphodiester linkages of natural RNA may be modified to include at least one of an nitrogen or sulfur heteroatom. Modifications in RNA structure may be tailored to allow specific genetic inhibition while avoiding a general response to dsRNA. Likewise, bases may be modified to block the activity of adenosine deaminase. The RNAi construct may be produced enzymatically or by partial/total organic synthesis, any modified ribonucleotide can be introduced by in vitro enzymatic or organic synthesis.

[0279] Methods of chemically modifying RNA molecules can be adapted for modifying RNAi constructs (see, for example, Heidenreich et al. (1997) *Nucleic Acids Res.*, 25:776-780; Wilson et al. (1994) *J Mol Recog* 7:89-98; Chen et al. (1995) *Nucleic Acids Res* 23:2661-2668; Hirschbein et al. (1997) *Antisense Nucleic Acid Drug Dev* 7:55-61). Merely to illustrate, the backbone of an RNAi construct can be modified with phosphorothioates, phosphoramidate, phosphodithioates, chimeric methylphosphonate-phosphodiesters, peptide nucleic acids, 5-propynyl-pyrimidine containing oligomers or sugar modifications (e.g., 2'-substituted ribonucleosides, a-configuration).

[0280] The double-stranded structure may be formed by a single self-complementary RNA strand or two complementary RNA strands. RNA duplex formation may be initiated either inside or outside the cell. The RNA may be introduced in an amount which allows delivery of at least one copy per cell. Higher doses (e.g., at least 5, 10, 100, 500 or 1000 copies per cell) of double-stranded material may yield more effective inhibition, while lower doses may also be useful for specific applications. Inhibition is sequence-specific in that nucleotide sequences corresponding to the duplex region of the RNA are targeted for genetic inhibition.

[0281] In certain embodiments, the subject RNAi constructs are "small interfering RNAs" or "siRNAs." These nucleic acids are around 19-30 nucleotides in length, and even more preferably 21-23 nucleotides in length, e.g., corresponding in length to the fragments generated by nuclease "dicing" of longer double-stranded RNAs. The siRNAs are understood to recruit nuclease complexes and guide the complexes to the target mRNA by pairing to the specific sequences. As a result, the target mRNA is degraded by the

nucleases in the protein complex. In a particular embodiment, the 21-23 nucleotides siRNA molecules comprise a 3' hydroxyl group.

[0282] The siRNA molecules of the present invention can be obtained using a number of techniques known to those of skill in the art. For example, the siRNA can be chemically synthesized or recombinantly produced using methods known in the art. For example, short sense and antisense RNA oligomers can be synthesized and annealed to form double-stranded RNA structures with 2-nucleotide overhangs at each end (Caplen, et al. (2001) *Proc Natl Acad Sci USA*, 98:9742-9747; Elbashir, et al. (2001) *EMBO J*, 20:6877-88). These double-stranded siRNA structures can then be directly introduced to cells, either by passive uptake or a delivery system of choice, such as described below.

[0283] In certain embodiments, the siRNA constructs can be generated by processing of longer double-stranded RNAs, for example, in the presence of the enzyme dicer. In one embodiment, the *Drosophila* in vitro system is used. In this embodiment, dsRNA is combined with a soluble extract derived from *Drosophila* embryo, thereby producing a combination. The combination is maintained under conditions in which the dsRNA is processed to RNA molecules of about 21 to about 23 nucleotides.

[0284] The siRNA molecules can be purified using a number of techniques known to those of skill in the art. For example, gel electrophoresis can be used to purify siRNAs. Alternatively, non-denaturing methods, such as non-denaturing column chromatography, can be used to purify the siRNA. In addition, chromatography (e.g., size exclusion chromatography), glycerol gradient centrifugation, affinity purification with antibody can be used to purify siRNAs.

[0285] In certain preferred embodiments, at least one strand of the siRNA molecules has a 3' overhang from about 1 to about 6 nucleotides in length, though may be from 2 to 4 nucleotides in length. More preferably, the 3' overhangs are 1-3 nucleotides in length. In certain embodiments, one strand having a 3' overhang and the other strand being blunt-ended or also having an overhang. The length of the overhangs may be the same or different for each strand. In order to further enhance the stability of the siRNA, the 3' overhangs can be stabilized against degradation. In one embodiment, the RNA is stabilized by including purine nucleotides, such as adenine or guanosine nucleotides. Alternatively, substitution of pyrimidine nucleotides by modified analogues, e.g., substitution of uridine nucleotide 3' overhangs by 2'-deoxythymidine is tolerated and does not affect the efficiency of RNAi. The absence of a 2' hydroxyl enhances the nuclease resistance of the overhang in tissue culture medium and may be beneficial in vivo.

[0286] In other embodiments, the RNAi construct is in the form of a long double-stranded RNA. In certain embodiments, the RNAi construct is at least 25, 50, 100, 200, 300 or 400 bases. In certain embodiments, the RNAi construct is 400-800 bases in length. The double-stranded RNAs are digested intracellularly, e.g., to produce siRNA sequences in the cell. However, use of long double-stranded RNAs in vivo is not always practical, presumably because of deleterious effects which may be caused by the sequence-independent dsRNA response. In such embodiments, the use of local delivery systems and/or agents which reduce the effects of interferon or PKR are preferred.

[0287] In certain embodiments, the RNAi construct is in the form of a hairpin structure (named as hairpin RNA). The

hairpin RNAs can be synthesized exogenously or can be formed by transcribing from RNA polymerase III promoters *in vivo*. Examples of making and using such hairpin RNAs for gene silencing in mammalian cells are described in, for example, Paddison et al., *Genes Dev.* 2002, 16:948-58; McCaffrey et al., *Nature*, 2002, 418:38-9; McManus et al., *RNA*, 2002, 8:842-50; Yu et al., *Proc Natl Acad Sci USA*, 2002, 99:6047-52). Preferably, such hairpin RNAs are engineered in cells or in an animal to ensure continuous and stable suppression of a desired gene. It is known in the art that siRNAs can be produced by processing a hairpin RNA in the cell.

[0288] In yet other embodiments, a plasmid is used to deliver the double-stranded RNA, e.g., as a transcriptional product. In such embodiments, the plasmid is designed to include a “coding sequence” for each of the sense and anti-sense strands of the RNAi construct. The coding sequences can be the same sequence, e.g., flanked by inverted promoters, or can be two separate sequences each under transcriptional control of separate promoters. After the coding sequence is transcribed, the complementary RNA transcripts base-pair to form the double-stranded RNA.

[0289] PCT application WO01/77350 describes an exemplary vector for bi-directional transcription of a transgene to yield both sense and antisense RNA transcripts of the same transgene in a eukaryotic cell. Accordingly, in certain embodiments, the present invention provides a recombinant vector having the following unique characteristics: it comprises a viral replicon having two overlapping transcription units arranged in an opposing orientation and flanking a transgene for an RNAi construct of interest, wherein the two overlapping transcription units yield both sense and antisense RNA transcripts from the same transgene fragment in a host cell.

[0290] RNAi constructs can comprise either long stretches of double stranded RNA identical or substantially identical to the target nucleic acid sequence or short stretches of double stranded RNA identical to substantially identical to only a region of the target nucleic acid sequence. Exemplary methods of making and delivering either long or short RNAi constructs can be found, for example, in WO01/68836 and WO01/75164.

[0291] Exemplary RNAi constructs that specifically recognize a particular gene, or a particular family of genes can be selected using methodology outlined in detail above with respect to the selection of antisense oligonucleotide. Similarly, methods of delivery RNAi constructs include the methods for delivery antisense oligonucleotides outlined in detail above.

[0292] Peptidomimetics: In other embodiments, the invention contemplates that the agent is a peptidomimetic. Peptidomimetics are compounds based on, or derived from, peptides and proteins. Peptidomimetics can be obtained by structural modification of the amino acid sequence of a known protein using unnatural amino acids, conformational restraints, isosteric replacement, and the like. The subject peptidomimetics constitute the continuum of structural space between peptides and non-peptide synthetic structures.

[0293] Exemplary peptidomimetics can have such attributes as being non-hydrolyzable (e.g., increased stability against proteases or other physiological conditions which degrade the corresponding peptide), having increased specificity and/or potency, and having increased cell permeability for intracellular localization. For illustrative purposes, pep-

tide analogs of the present invention can be generated using, for example, benzodiazepines (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988, p 123), C-7 mimics (Huffman et al. in *Peptides: Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988, p. 105), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29:295; and Ewenson et al. in *Peptides. Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, Ill., 1985), beta-turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26:647; and Sato et al. (1986) *J Chem Soc Perkin Trans 1*:1231), beta-aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126:419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134:71), diaminoketones (Natarajan et al. (1984) *Biochem Biophys Res Commun* 124:141), and methyleneamino-modified (Roark et al. in *Peptides: Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988, p 134). Also, see generally, Session III: Analytic and synthetic methods, in *Peptides: Chemistry and Biology*, G. R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988)

[0294] In addition to a variety of sidechain replacements which can be carried out to generate the subject peptidomimetics, the present invention specifically contemplates the use of conformationally restrained mimics of peptide secondary structure. Numerous surrogates have been developed for the amide bond of peptides. Frequently exploited surrogates for the amide bond include the following groups (i) trans-olefins, (ii) fluoroalkene, (iii) methyleneamino, (iv) phosphonamides, and (v) sulfonamides.

[0295] Additionally, peptidomimetics based on more substantial modifications of the backbone of a peptide can be used. Peptidomimetics which fall in this category include (i) retro-inverso analogs, and (ii) N-alkyl glycine analogs (so-called peptoids).

[0296] Furthermore, the methods of combinatorial chemistry are being brought to bear, e.g., PCT publication WO 99/48897, on the development of new peptidomimetics. For example, one embodiment of a so-called “peptide morphing” strategy focuses on the random generation of a library of peptide analogs that comprise a wide range of peptide bond substitutes.

[0297] In an exemplary embodiment, the peptidomimetic can be derived as a retro-inverso analog of the peptide. Retro-inverso analogs can be made according to the methods known in the art, such as that described by the Sisto et al. U.S. Pat. No. 4,522,752. As a general guide, sites which are most susceptible to proteolysis are typically altered, with less susceptible amide linkages being optional for mimetic switching. The final product, or intermediates thereof, can be purified by HPLC.

[0298] In another illustrative embodiment, the peptidomimetic can be derived as a retro-enatio analog of a peptide. Retro-enantio analogs such as this can be synthesized using commercially available D-amino acids (or analogs thereof) and standard solid- or solution-phase peptide-synthesis techniques. For example, in a preferred solid-phase synthesis method, a suitably amino-protected (t-butyloxycarbonyl, Boc) residue (or analog thereof) is covalently bound to a solid support such as chloromethyl resin. The resin is washed with

dichloromethane (DCM), and the BOC protecting group removed by treatment with TFA in DCM. The resin is washed and neutralized, and the next Boc-protected D-amino acid is introduced by coupling with diisopropylcarbodiimide. The resin is again washed, and the cycle repeated for each of the remaining amino acids in turn. When synthesis of the protected retro-enantio peptide is complete, the protecting groups are removed and the peptide cleaved from the solid support by treatment with hydrofluoric acid/anisole/dimethyl sulfide/thioanisole. The final product is purified by HPLC to yield the pure retro-enantio analog.

[0299] In still another illustrative embodiment, trans-olefin derivatives can be made for any of the subject polypeptides. A trans olefin analog can be synthesized according to the method of Y. K. Shue et al. (1987) *Tetrahedron Letters* 28:3225 and also according to other methods known in the art. It will be appreciated that variations in the cited procedure, or other procedures available, may be necessary according to the nature of the reagent used.

[0300] It is further possible to couple the pseudodipeptides synthesized by the above method to other pseudodipeptides, to make peptide analogs with several olefinic functionalities in place of amide functionalities.

[0301] Still another class of peptidomimetic derivatives include phosphonate derivatives. The synthesis of such phosphonate derivatives can be adapted from known synthesis schemes. See, for example, Loots et al. in *Peptides: Chemistry and Biology*, (Escom Science Publishers, Leiden, 1988, p. 118); Petrillo et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium, Pierce Chemical Co. Rockland, Ill., 1985).

[0302] Many other peptidomimetic structures are known in the art and can be readily adapted for use in designing peptidomimetics. To illustrate, the peptidomimetic may incorporate the 1-azabicyclo[4.3.0]nonane surrogate (see Kim et al. (1997) *J. Org. Chem.* 62:2847), or an N-acyl piperazic acid (see Xi et al. (1998) *J. Am. Chem. Soc.* 120:80), or a 2-substituted piperazine moiety as a constrained amino acid analogue (see Williams et al. (1996) *J. Med. Chem.* 39:1345-1348). In still other embodiments, certain amino acid residues can be replaced with aryl and bi-aryl moieties, e.g., monocyclic or bicyclic aromatic or heteroaromatic nucleus, or a biaromatic, aromatic, heteroaromatic, or biheteroaromatic nucleus.

[0303] Small organic or inorganic molecules: In certain embodiments, the compound is a small organic or inorganic molecule. Small organic or inorganic molecules can agonize or antagonize the function of a particular protein or class of proteins. By small organic or inorganic molecule is meant a carbon contain molecule having a molecular weight less than 5000 amu, preferably less than 2500 amu, more preferably less than 1500 amu, and even more preferably less than 750, 500, or 250 amu.

[0304] Small organic or inorganic molecules can be readily identified by screening libraries of organic molecules and/or chemical compounds to identify those compounds that have a desired function. Alternatively, single compounds or small numbers of candidate compounds can be screened individual or in combination. In certain embodiments, the small molecule (e.g., an inorganic or organic molecule) is a non-peptidyl compound containing two or fewer, one or fewer, or no peptide and/or saccharide linkages.

[0305] The foregoing are illustrative examples of classes of compounds that can be used in the various methods and

screening assays of the present invention. One of skill in the art can select amongst available delivery methods to deliver the compound to the particular cells in vitro or in vivo. By way of example, many compounds readily transit epidermal barriers and other biological membranes. To administer such compounds to cells or to an animal, the compound can simply be dissolved and added to the fluid in which the cells or animal is cultured. Alternatively, the compound can be dissolved and added to the animals food or drinking water. In another alternative, the compound can be administered to the animal via local or systemic injection.

[0306] Certain compounds do not as readily transit epidermal barriers and biological membranes, and thus additional techniques have been adapted to administer such compounds to cells, tissues, and organisms. For example, RNAi constructs are often administered to animals by addition to their food or drinking water. Numerous types of nucleic acids are delivered via viral or plasmid-based expression vectors. Polypeptide-based compounds that do not readily transit membrane or that are not actively transported into cells via receptor-mediated mechanisms can be administered along with carriers that facilitate transit into cells and tissues. The foregoing exemplary administration methods are well known in the art and can be selected based on the compounds and organisms being employed in the particular screening assays or methods of use.

[0307] Whether compounds are being administered as part of a screening assay or as part of a method for modulating cells behavior, compounds can be administered alone or as pharmaceutical formulations. Exemplary pharmaceutical compositions are formulated for administration to cells or animals. In certain embodiments, the compound included in the pharmaceutical preparation may be active itself, or may be a prodrug, e.g., capable of being converted to an active compound in a physiological setting. In certain embodiments the subject compounds may be simply dissolved or suspended water, for example, in sterile water. In certain embodiments, the pharmaceutical preparation is non-pyrogenic, i.e., does not elevate the body temperature of an animal.

[0308] The phrase "pharmaceutically acceptable carrier" as used herein means a pharmaceutically acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting a subject compound. Each carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the patient. Some examples of materials which can serve as pharmaceutically acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) phosphate buffer solutions; and (21) other non-toxic compatible substances employed in pharmaceutical formulations.

[0309] As set out above, in certain embodiments the agents may contain a basic functional group, such as amino or alkylamino, and are, thus, capable of forming pharmaceutically acceptable salts with pharmaceutically acceptable acids. The term "pharmaceutically acceptable salts" in this respect, refers to the relatively non-toxic, inorganic and organic acid addition salts of compounds of the present invention. These salts can be prepared in situ during the final isolation and purification of the compounds of the invention, or by separately reacting a purified compound of the invention in its free base form with a suitable organic or inorganic acid, and isolating the salt thus formed. Representative salts include the hydrobromide, hydrochloride, sulfate, bisulfate, phosphate, nitrate, acetate, valerate, oleate, palmitate, stearate, laurate, benzoate, lactate, phosphate, tosylate, citrate, maleate, fumarate, succinate, tartrate, naphthylate, mesylate, glucoheptonate, lactobionate, and laurylsulphonate salts and the like. (See, for example, Berge et al. (1977) "Pharmaceutical Salts", *J. Pharm. Sci.* 66:1-19)

[0310] The pharmaceutically acceptable salts of the subject compounds include the conventional nontoxic salts or quaternary ammonium salts of the compounds, e.g., from non-toxic organic or inorganic acids. For example, such conventional nontoxic salts include those derived from inorganic acids such as hydrochloride, hydrobromic, sulfuric, sulfamic, phosphoric, nitric, and the like; and the salts prepared from organic acids such as acetic, propionic, succinic, glycolic, stearic, lactic, malic, tartaric, citric, ascorbic, palmitic, maleic, hydroxymaleic, phenylacetic, glutamic, benzoic, salicyclic, sulfanilic, 2-acetoxybenzoic, fumaric, toluene-sulfonic, methanesulfonic, ethane disulfonic, oxalic, isothionic, and the like.

[0311] In other cases, the compounds of the present invention may contain one or more acidic functional groups and, thus, are capable of forming pharmaceutically acceptable salts with pharmaceutically acceptable bases. The term "pharmaceutically acceptable salts" in these instances refers to the relatively non-toxic, inorganic and organic base addition salts of compounds of the present invention. These salts can likewise be prepared in situ during the final isolation and purification of the compounds, or by separately reacting the purified compound in its free acid form with a suitable base, such as the hydroxide, carbonate or bicarbonate of a pharmaceutically acceptable metal cation, with ammonia, or with a pharmaceutically acceptable organic primary, secondary or tertiary amine. Representative alkali or alkaline earth salts include the lithium, sodium, potassium, calcium, magnesium, and aluminum salts and the like. Representative organic amines useful for the formation of base addition salts include ethylamine, diethylamine, ethylenediamine, ethanolamine, diethanolamine, piperazine and the like. (See, for example, Berge et al., *supra*)

[0312] Wetting agents, emulsifiers and lubricants, such as sodium lauryl sulfate and magnesium stearate, as well as coloring agents, release agents, coating agents, sweetening, flavoring and perfuming agents, preservatives and antioxidants can also be present in the compositions.

[0313] Examples of pharmaceutically acceptable antioxidants include: (1) water soluble antioxidants, such as ascorbic acid, cysteine hydrochloride, sodium bisulfite, sodium metabisulfite, sodium sulfite and the like; (2) oil-soluble antioxidants, such as ascorbyl palmitate, butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), lecithin, propyl gallate, alpha-tocopherol, and the like; and (3) metal chelat-

ing agents, such as citric acid, ethylenediamine tetraacetic acid (EDTA), sorbitol, tartaric acid, phosphoric acid, and the like.

[0314] In certain aspects and embodiments, the present invention provides screening assays. In certain embodiments, the assays are screening assays which utilize compounds as a tool to identify and/or characterize a role for an ion transporter protein or a class of ion transporter protein during a particular biological process. Once identified, the compounds, candidate ion transporter proteins, or candidate class of ion transporter proteins can be further studied in other animals or in cell-based or cell-free assays *in vitro*.

[0315] In certain aspects and embodiments, the present invention provides screening assays to identify and/or characterize compounds (either known or novel) that modulate ion flux and/or membrane potential in a population of cells. Such compounds can be used to modulate cell behavior, for example to modulate cell dedifferentiation and/or regeneration, *in vitro* or *in vivo*. Identified compounds may be further characterized. Identified compounds may be useful for research directed to the further study of a particular behavioral, anatomical, or morphological process. Identified compounds may be useful for research directed to the further study of a particular ion transporter protein or class of ion transporter proteins. Furthermore, identified compounds may be useful in the development of a pharmaceutical, or even as a pharmaceutical product. Accordingly, the present invention provides for compounds and pharmaceutical compounds identified and/or characterized by any of the methods of the invention.

[0316] In certain embodiments, the identified compounds may be useful in the development of a pharmaceutical product. Further testing of a possible pharmaceutical product may involve administration to animals and may additionally involve study of the preferable route of administration. Thus, another aspect of the present invention provides pharmaceutically acceptable compositions comprising an effective amount of one or more of the compounds described above, formulated together with one or more pharmaceutically acceptable carriers (additives) and/or diluents. The pharmaceutical compositions of the present invention may be specially formulated for administration in solid or liquid form, including those adapted for the following: (1) oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets, boluses, powders, granules, pastes for application to the tongue; (2) parenteral administration, for example, by subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution or suspension; (3) topical application, for example, as a cream, ointment or spray applied to the skin; (4) intravaginally or intrarectally, for example, as a pessary, cream or foam; or (5) ophthalmic administration, for example, for administration following injury or damage to the retina. However, in certain embodiments the subject compounds may be simply dissolved or suspended in sterile water. In certain embodiments, the pharmaceutical preparation is non-pyrogenic, i.e., does not elevate the body temperature of a human or animal patient.

(iii) Cells and Animals

[0317] As outlined throughout in reference to particular methods of the invention, the subject methods and screening assays can be conducted *in vivo* or *in vitro* in cells derived from or resident in virtually any organism. Exemplary organ-

isms include, but are not limited to, plants, prokaryotes (e.g., bacteria), and fungi. Further exemplary organisms include, but are not limited to, chordates, hemichordates, protochordates, and invertebrates.

[0318] The foregoing methods can be conducted in cells in culture, in tissue samples maintained ex vivo, or in animals. When the method is conducted using cells in culture, the invention contemplates using cells derived from any organism, tissue, or stage of development. Furthermore, the invention contemplates that the cells may be primary cultures of cells, or transformed cell lines, and that the cells can either be wild type cells or cells containing one or more mutations. Mutant cells or cell lines may be models of a particular disease or injury, or may be derived from animals having a specific disease or injury (e.g. cancer cells harvested from an animal).

[0319] Cells may be derived from (e.g., derived from and cultured in vitro as populations of cells or tissues) or reside in (cells resident in a whole animal or portion of a whole animal) any of a number of animal species. Exemplary animals include, but are not limited to, flatworms, amphibians, fish, reptiles, birds, or mammals. Suitable flatworms include planarian. Suitable amphibians include *Xenopus laevis*, *Xenopus tropicalis*, and other species of frog. Suitable birds include chickens, as well as other birds commonly used or maintained in a laboratory setting. Suitable mammals include mice, rats, hamsters, goats, sheep, pigs, cows, dogs, cats, rabbits, non-human primates, and humans.

[0320] Regardless of the species of cells or animal selected, the invention contemplates that cells may be derived from or reside in an animal of virtually any stage of development. For example, the cells may be derived from or reside in an embryonic, larval, fetal, juvenile, or adult organism. The decision of whether to conduct a particular screen in cells, tissues, or animals, as well as the species and stage of development of the selected cells or animals can be readily made by one of skill in the art. The skilled artisan can select the approach, conditions, and system based on their expertise, resources, and the particular biological process they are investigating.

[0321] To further illustrate, in one embodiment, the foregoing methods are conducted using cells derived from or resident in a nematode. There are over 10,000 known nematode species. These include parasitic nematodes (e.g., nematodes that are parasitic to humans, non-human animals, or plants). Exemplary parasitic nematodes include, but are not limited to, whipworms, *Ascaris*, hookworms, filarial worms, and root knot nematodes. *C. elegans* is perhaps the most well known and thoroughly studied nematode, and the invention contemplates using *C. elegans* or other nematodes.

[0322] In another embodiment, the foregoing methods are conducted in cells derived from or resident in a fish or amphibian species. Zebrafish (e.g., adult zebrafish and developing, e.g., embryonic fish) are a particular example of a fish well suited for study. Zebrafish are an extensively used developmental system, and genetic, cell biological, and molecular biological reagents and methods are well known and available. Additionally, numerous chemical and radiation-based screens have produced large numbers of mutant zebrafish that can also be used for study.

[0323] *Xenopus laevis* and *Xenopus tropicalis* (e.g., adult, embryonic, tadpole, etc. stage animals) are particular examples of amphibians well suited for study. Both species are used extensively, and well developed reagents exist. For example, the availability of these molecular reagents facili-

tates screening assays based on changes in gene or protein expression, either instead of or in addition to screening assays based on morphological criteria. Additionally, *Xenopus tropicalis* is a genetically tractable model organism, and mutants have been and continue to be generated and characterized. *Xenopus* cells and whole organisms are excellent systems for screening assays. The cells of early *Xenopus* embryos are relatively large, and thus easily manipulated, injected, and used for electrophysiological recording. Eggs and embryos can be collected in very large numbers. This allows high-throughput screening, and facilitates biochemical, pharmacological, and statistical analyses.

[0324] In another embodiment, the foregoing methods are conducted in cells derived from or resident in a flatworm. Exemplary flatworms are the free-living (e.g., non-parasitic) flatworm planaria. Planaria are in the phylum Platyhelminthes and the class Turbellaria. There are numerous species of planaria, any of which can be readily used. Planaria exhibit much of the complexity of vertebrate systems: a well-differentiated nervous system, intestine, eyes, brain, three tissue layers, and bilateral symmetry. Planaria represent a critical breakthrough in the evolution of the animal body plan and are thought to very closely resemble the proto-bilaterian ancestor. It is the first organism to have both bilateral symmetry and encephalization, making it capable of detecting environmental stimuli quicker and more efficiently than the lower metazoans. Despite a simplistic appearance and evolutionary position, planaria possess a well-developed nervous system with true synaptic transmission and have what can be considered the first animal "brain" (Samat and Netsky (1985) Can J Neurol Sci. 12(4): 296-302). They have also developed sensory capabilities for the detection of light (Brown and Park (1975) Int J Chronobiol. 3(1):57-62; Brown et al., 1968), chemical gradients (Mason (1975) Anim Behav. May; 23(2): 460-9; Miyamoto and Shimozawa (1985) Zoological Science (Tokyo) 2: 389-396), vibration (Fulgheri and Messeri (1973) Boll Soc Ital Biol Sper. 49(20): 1141-5), electric fields (Brown and Ogden (1968) J Gen Physiol. 51(2):255-60), magnetic fields (Brown and Chow (1975) Physiological Zoology 48: 168-176; Brown (1966) Nature 209: 533-5), and weak γ -radiation (Brown and Park (1964) Nature 202: 469-471).

[0325] Like many of the other organisms described above, planaria are well suited for screening because of their small size. Furthermore, they are easy to raise and to subject to a multitude of reagents and manipulations. Their consistent, flat shape and active behaviors make it simple to observe the results of any behavioral or morphological perturbation. Moreover, evolutionarily, they are very similar to the ancestor of the *bilateria* clade, and have high relevance to human medicine and physiology both structurally and physiologically (Best and Morita (1982) Teratog Carcinog Mutagen. 2(3-4): 277-91; Samat and Netsky (1985) Can J Neurol Sci. 12(4): 296-302). Planaria offer an excellent combination of experimental tractability and sufficient complexity for asking a number of fascinating questions about basic functions of living systems (Eisenstein (1997) Behavioral Brain Research 82: 121-132). Crucially, as a model system, planaria are quickly acquiring a powerful set of molecular biological reagents and techniques, enabling genetic and cell-biological investigations into its structure and function (Agata et al., 2003; Alvarado et al., 2002; Cebria et al. (2002) Nature 419: 620-4).

[0326] Planaria have exceptional regenerative capacity. A bisected flatworm readily regenerates. Thus planaria, either

whole animals or fragments, serve as an excellent model system in which to study the implications of ion flux on cell dedifferentiation and regeneration, as well as on progenitor cell identification and characterization. In addition to planaria, other model systems have enhanced regenerative capacity, and these systems are especially well suited for studies of ion flux on cell dedifferentiation and regeneration. By way of example, fish and amphibian species may be especially useful as model systems in such studies.

[0327] The invention contemplates the use of any of the foregoing animals, as well as plants, prokaryotes, and fungi. Each of these has numerous characteristics that make them suitable for particular screening assays. The appropriate model organism can be readily selected based on the particular assays being conducted, as well as space and resource constraints. An investigator can readily determine whether to conduct the assays in cells cultured in vitro, tissue explants cultured in vitro, or in whole animals or animal fragments. Furthermore, the appropriate developmental stage can be readily selected. Exemplary developmental stages include, but are not limited to, embryonic stages, tadpole stages, larval stages, juvenile stages, and adult stages. Furthermore, the invention contemplates studying whole animals, as well as animal fragments. An exemplary animal fragment is a bisected or trisected organism. In one embodiment, the animal fragment is a bisected or trisected planarian. In another embodiment, the animal fragment is formed by fission of a whole animal. Additionally, the invention contemplates the use of wild type or mutant animals. In one embodiment, the animal is a wild type embryonic, tadpole, larval, fetal, juvenile, or adult stage animal. In another embodiment, the animal is a mutant embryonic, tadpole, larval, fetal, juvenile, or adult stage animal. In still another embodiment, the animal includes an injury. In yet another embodiment, the animal is a model of regeneration (e.g., an animal or tissue in an animal that endogenously possesses robust regenerative capacity).

[0328] It is often informative to use more than one species, both to overcome difficulties inherent in one type of model system, as well as to gain insight into the evolutionary biology of ion flux related phenomena.

[0329] Many of the foregoing animals are particularly attractive due to the availability of additional experimental tools and reagents to facilitate assay methods, as well as further study. For example, cell biological, genetic, and/or molecular reagents exist for many of these model organisms. To illustrate, the availability of fluorescent reagents to label living or non-living animals and cells may facilitate further study and analysis. Additional cell and molecular tools including reagents for RNA analysis (e.g., Northern blot hybridization, RT-PCR, RNase protection, in situ hybridization, GeneChip analysis) and/or protein analysis (e.g., immunohistochemistry, Western blot analysis) may be useful.

[0330] Regardless of the particular model organism selected, and regardless of whether the subject methods are conducted in vitro or in vivo, cells or animals may be of any developmental stage including, but not limited to, embryonic, fetal, larval, tadpole, juvenile, and adult stage cells or organisms. One of skill in the art can select the proper animal and developmental stage depending on the particular assay being conducted, the particular compounds being assessed, and the particular developmental or behavioral process being investigated. Furthermore, one of skill in the art can select the appropriate animal and developmental stage based on the research interests of the investigator, time, and cost consid-

erations, as well as the availability of other complementary research reagents. Additionally, even when whole organisms or large fragments of whole organisms are used, one of skill in the art may choose to examine a particular biological process in only a portion of the whole organism or fragment.

[0331] In one embodiment, the animal or tissue (including whole animals, injured animal, fragments, or cell derived therefrom) is selected based on its robust regenerative ability. Cells, tissues, or animals with an enhanced regenerative ability may be useful in methods for identifying and characterizing a role for ion transporter proteins, ion flux, membrane potential, and/or pH in dedifferentiation and regeneration. Exemplary animals and systems with enhanced regenerative capacity include, but are not limited to, planaria, the zebrafish tail, the amphibian (e.g., *Xenopus*) tail, and the amphibian limb. An understanding of how regeneration is modulated in any of these systems can be used to increase/stimulate regenerative capacity in organisms and systems whose endogenous regenerative capacity is less robust. In another embodiment, the animal or tissue is selected for screening and study specifically because its endogenous regenerative capacity is not robust. Such systems include any cells or tissues derived from organisms, such as mammals, whose endogenous regenerative capacity is not robust. Such systems also include endogenously non-regenerating cells or tissues derived from particular regions of otherwise robustly regenerative organisms.

[0332] In one embodiment, the animal/organism is a protostome. Protostomes possess a hollow dorsal nerve cord, gill slits, and a notochord. Exemplary protostomes include tunicata (e.g., sea squirts, etc.) and cephalochordate (e.g., amphioxus). Exemplary amphioxus include, but are not limited to *Ciona intestinalis* and *Branchiostoma floridae* (Holland and Gibson-Brown (2003) BioEssays 25: 528-532; Gostling and Shimeld (2003) Evolution and Development 5: 136; Dehal et al. (2002) Science 298: 2157-2167; Nishiyama et al. (1972) Tohoku J Exp Med 107: 95-96; Ogasawara et al. (2002) Develop Genes Evol 212: 173-185; Pope and Rowley (2002) J Exp Biology 205: 1577-1583).

[0333] In another embodiment, the animal/organism is a hemichordate. Exemplary hemichordates include acorn worms (Tagawa et al. (2001) Evol and Develop 3: 443).

[0334] In another embodiment, the animal/organism is a nematode. There are over 10,000 known nematode species. These include parasitic nematodes (e.g., nematodes that are parasitic to humans, non-human animals, or plants). Exemplary parasitic nematodes include, but are not limited to, whipworms, *Ascaris*, hookworms, filarial worms, and root knot nematodes.

[0335] *C. elegans* is perhaps the most well known and thoroughly studied nematode, and the invention contemplates using *C. elegans* or other nematodes. Although *C. elegans* is considered a soil nematode, methods for culturing *C. elegans* in various quantities of liquid media (e.g., in a fluid) are well developed. See, <http://elegans.swmed.edu/>. Accordingly, the methods and apparatuses of the invention for conducting assays in aquatic animals can be readily used to conduct assays in *C. elegans*.

[0336] In another embodiment, the animal is a fish or amphibian. Exemplary amphibians include frog (e.g., species of *Xenopus*) and salamanders (e.g., species of *Axolotl*).

[0337] In another embodiment, the animal is a flatworm. Exemplary flatworms are the free-living (e.g., non-parasitic) flatworm planaria. Planaria are in the phylum Platyhelmin-

thes and the class Turbellaria. There are numerous species of planaria, any of which can be readily used.

[0338] In another embodiment, the organism is a mammal such as a mouse, rat, rabbit, pig, cow, dog, cat, non-human primate, or human.

[0339] The invention contemplates the use of any of the foregoing animals. Each of these has numerous characteristics that make them suitable for particular screening assays or for particular methods of promoting/inhibiting dedifferentiation and/or regeneration. The appropriate model organism can be readily selected based on the particular assays being conducted, as well as space and resource constraints. Furthermore, the appropriate developmental stage can be readily selected. Exemplary developmental stages include, but are not limited to, embryonic stages, fetal stages, tadpole stages, larval stages, juvenile stages, and adult stages. In certain embodiments, the animal is chosen due to its optical accessibility. Furthermore, the invention contemplates studying whole animals, animal fragments, or animals inflicted with an injury. An exemplary animal fragment is a bisected or trisected organism. In one embodiment, the animal fragment is a bisected or trisected planarian. In another embodiment, the animal fragment is formed by fission of a whole animal. Additionally, the invention contemplates the use of wild type or mutant animals. In one embodiment, the animal is a wild type embryonic, tadpole, larval, fetal, juvenile, or adult stage animal. In another embodiment, the animal is a mutant embryonic, tadpole, larval, fetal, juvenile, or adult stage animal.

[0340] In certain embodiments, it may be desirable to conduct an assay, for example an assay to identify and/or characterize a compound that modulates a particular developmental process, in a relatively simple system. Identified compounds or candidate ion transporter proteins can later be analyzed in higher organisms including mice, rats, non-human primates, and humans.

[0341] In certain other embodiments, it may be desirable to conduct an assay in parallel using different populations of cells. For example, screening assays can be conducted in parallel using cells derived from or resident in different organisms. Alternatively, screening assays can be conducted in parallel using cells of varying developmental stages derived from or resident in the same organism. In still another embodiment, screening assays can be conducted in parallel using cells of different developmental lineages (e.g., different cell or tissue types) derived from or resident in the same model organism. In this embodiment, the cells of differing developmental lineages can be of the same or varying developmental stages.

[0342] Depending on the particular model system and biological process chosen (e.g., organism, cell type, developmental stage, etc) for study or manipulation, one of skill in the art can select the appropriate culture conditions and methods for monitoring changes in the model system. For example, certain phenotype changes can be observed and monitored based on visual inspection with either the aided or unaided eye. Other phenotypic changes can be observed using molecular, cell biological, or biophysical reagents available in the art. For example, changes in the expression of one or more molecular markers can be assessed using known techniques including, but not limited to, RT-PCR, *in situ* hybridization, Northern blot analysis, Western blot analysis, immunocytochemistry, immunohistochemistry, and GeneChip analysis. Further tools including, but not limited to, method of

detecting changes in cell proliferation, cell death, cell survival, membrane potential, intracellular pH, ion flux and the like can also be used to detect and assess phenotypic changes in cells or organisms.

(iv) Exemplary Diseases and Injuries

[0343] As outlined above, the present invention provides methods for identifying ion transporter proteins and classes of ion transporter proteins that mediate ion flux and/or membrane potential. Identified transporters, as well as compounds that modulate (e.g., inhibit or promote) the activity of those transporters may be useful for modulating a biological process *in vitro* or *in vivo*. In certain embodiments, one or more ion transporter proteins or compounds that modulate the expression and/or activity of ion transporter proteins may be useful in modulating dedifferentiation and/or regeneration. The invention contemplates the use of a single compound to modulate a single ion transporter protein or class of ion transporter proteins. The invention further contemplates the use of multiple compounds (2, 3, 4, etc) to modulate a single ion transporter protein or a class of ion transporter proteins. The invention further contemplates the use of multiple compounds (2, 3, 4, etc) to modulate a multiple ion transporter proteins or multiple classes of ion transporter proteins.

[0344] Compounds, and pharmaceutical preparations thereof, that modulate dedifferentiation and/or regeneration may be useful in the treatment of injury or degenerative disease. Such compounds can be administered to a human or non-human patient in need of augmenting a regenerative response to disease or injury. Briefly, compounds that promote regeneration may be administered to promote the combination of proliferation, differentiation, and/or dedifferentiation needed to regenerate damaged, diseased, or injured tissue.

[0345] The invention contemplates the use of compounds individually or in combination. Suitable combinations include combinations of multiple compounds identified as promoting dedifferentiation and/or regeneration by modulating ion flux and/or membrane potential. Suitable combinations also include a compound that promotes dedifferentiation and/or regeneration by modulating ion flux and/or membrane potential along with one or more agents conventionally used in the treatment of the particular injury or degenerative disease.

[0346] Multiple agents may act additively or synergistically, and include combinations of agents that may show little or no effect when administered alone. Furthermore, the invention contemplates the use of agents in combination with known factors that influence proliferation, differentiation, or survival of a particular cell type. Still further, the invention contemplates the use of agents as part of a therapeutic regimen along with other surgical, radiological, chemical, homeopathic, or pharmacologic intervention appropriate for the particular cell type, disease or condition.

[0347] Agents which possess one of more of these characteristics may be useful in a therapeutic context. For example, injuries and diseases of the central and peripheral nervous system effect a tremendous number of people and exact a large financial and person toll. Injuries include traumatic injuries (i.e., breaks, blunt injury, burns, lacerations) to the brain or spinal cord, as well as other injuries to any region of the CNS or PNS including, but not limited to, injuries caused by bacterial infection, viral infection, cell damage following surgery, exposure to a toxic agent, cellular damage caused by

cancer or other proliferative disorder, ischemia, hypoxia, and the like. Currently, effective treatments for injuries of the CNS and PNS are limited, and individuals often experience long-term deficits consistent with the extent of injury, the location of the injury, and the types of cell that are effected.

[0348] In addition to injuries of the CNS and PNS, there are a wide variety of neurodegenerative diseases that effect particular regions and/or cell types of the CNS or PNS. These diseases are often progressive in nature, and individuals afflicted with many of these diseases have few treatment options at their disposal. Exemplary neurodegenerative diseases include, but are not limited to, Parkinson's disease, Huntington's disease, Alzheimer's disease, ALS, multiple sclerosis, stroke, macular degeneration, peripheral neuropathy, and diabetic neuropathy.

[0349] In certain embodiments, compounds can be administered to promote regeneration of mesodermal or endodermal cell and tissue types. Injuries and diseases of tissues derived from the mesoderm or endoderm include, but are not limited to, myocardial infarction, osteoarthritis, rheumatoid arthritis, diabetes, cirrhosis, polycystic kidney disease, inflammatory bowel disease, pancreatitis, Crohn's disease, cancer of any mesodermal or endodermal tissue (e.g., pancreatic cancer, Wilms tumor, soft cell carcinoma, bone cancer, breast cancer, prostate cancer, ovarian cancer, uterine cancer, liver cancer, colon cancer, etc.), and injuries to any mesodermal or endodermal tissue including breaks, tears, bruises, lacerations, burns, toxicity, bacterial infection, and viral infection.

[0350] Furthermore, agents identified by the methods of the present invention may be used to modulate cells of the blood and blood vessels. Exemplary agents can be used to modulate (promote or inhibit) angiogenesis. Inhibition of angiogenesis is of particular use in the treatment of many forms of cancers, as well as in conditions aggravated by excess angiogenesis such as macular degeneration. Promotion of angiogenesis is of particular use in the treatment of conditions caused or aggravated by decreased blood flow. Exemplary conditions include, but are not limited to, myocardial infarction, stroke, and ischemia. Additionally, agents identified by the methods of the present invention can be used to promote proliferation and differentiation of various cell types of the blood and can be used in the treatment of anemia, leukemia, and various immunodeficiencies.

[0351] For any of the foregoing, the application contemplates that agents may be administered alone, or may be administered in combination with other agents. Further, the application contemplates that agents identified according to the subject methods can be administered as part of a therapeutic regimen along with other treatments appropriate for the particular injury or disease being treated. For example, in the case of Parkinson's disease, a subject agent may be administered in combination with L-dopa or other Parkinson's disease medications, or in combination with a cell based neuronal transplantation therapy for Parkinson's disease. In the case of an injury to the brain or spinal cord, a subject agent may be administered in combination with physical therapy, hydrotherapy, massage therapy, and the like. In the case of peripheral neuropathy, as for example diabetic neuropathy, a subject agent may be administered in combination with insulin. In the case of myocardial infarction, the subject agent may be administered along with angioplasty, surgery, blood pressure medication, and/or as part of an exercise and diet regimen.

[0352] Physical injuries may result in cellular damage that ultimately limits the function of a particular cell or tissue. For example, physical injuries to cells in the CNS may limit the function of cells in the brain, spinal cord, or eye. Examples of physical injuries include, but are not limited to, crushing or severing of neuronal tissue, such as may occur following a fall, car accident, gun shot or stabbing wound, etc. Further examples of physical injuries include those caused by extremes in temperature such as burning, freezing, or exposure to rapid and large temperature shifts.

[0353] Physical injuries to mesodermal cell types include injuries to skeletal muscle, cardiac muscle, tendon, ligament, cartilage, bone, and the like. Examples of physical injuries include, but are not limited to, crushing, severing, breaking, bruising, and tearing of muscle tissue, bone or cartilage such as may occur following a fall, car accident, gun shot or stabbing wound, etc. Further examples of physical injuries include breaking, tearing, or bruising of muscle tissue, bone, cartilage, ligament, or tendon as may occur following a sports injury or due to aging. Further examples of physical injuries include those caused by extremes in temperature such as burning, freezing, or exposure to rapid and large temperature shifts.

[0354] Physical injuries to endodermal cell types include injuries to hepatocytes and pancreatic cell types. Examples of physical injuries include, but are not limited to, crushing, severing, and bruising, such as may occur following a fall, car accident, gun shot or stabbing wound, etc. Further examples of physical injuries include those caused by extremes in temperature such as burning, freezing, or exposure to rapid and large temperature shifts.

[0355] Further examples of an injury to any of the aforementioned cell types include those caused by infection such as by a bacterial or viral infection. Examples of bacterial or viral infections include, but are not limited to, meningitis, staph, HIV, hepatitis A, hepatitis B, hepatitis C, syphilis, human papilloma virus, strep, etc. However, one of skill in the art will recognize that many different types of bacteria or viruses may infect cells and cause injury.

[0356] Additionally, injury to a particular cell type may occur as a consequence or side effect of other treatments being used to relieve some condition in an individual. For example, cancer treatments (chemotherapy, radiation therapy, surgery) may cause significant damage to both cancerous and healthy cells. Surgery; implantation of intraluminal devices; the placement of implants, pacemakers, shunts; and the like can all result in cellular damage.

[0357] A wide range of neurodegenerative diseases cause extensive cell damage (i.e., injury) to cells of the CNS and PNS. Accordingly, neurodegenerative diseases are candidates for treatment using the described agents. Administration of a subject agent can promote neuronal regeneration in the CNS or PNS of a patient with a neurodegenerative disease, and the promotion of neuronal regeneration can ameliorate, at least in part, symptoms of the disease. Agents may be administered individually, in combination with other agents of the invention, or as part of a treatment regimen appropriate for the specific condition being treated. The following are illustrative examples of neurodegenerative conditions which can be treated using the subject agents.

[0358] Parkinson's disease is the result of the destruction of dopamine-producing neurons of the substantia nigra, and results in the degeneration of axons in the caudate nucleus and the putamen degenerate. Although therapies such as L-dopa

exist to try to ameliorate the symptoms of Parkinson's disease, to date we are unaware of treatments which either prevent the degeneration of axons and/or increase neuronal regeneration. Administration of agents which promote neuronal regeneration can help to ameliorate at least certain symptoms of Parkinson's disease including rigidity, tremor, bradykinesia, poor balance and walking problems.

[0359] Alzheimer's disease, a debilitating disease characterized by amyloid plaques and neurofibrillary tangles, results in a loss of nerve cells in areas of the brain that are vital to memory and other mental abilities. There also are lower levels of chemicals in the brain that carry complex messages back and forth between nerve cells. Alzheimer's disease disrupts normal thinking and memory. The incidence of Alzheimer's disease will only increase as the average life expectancy continues to rise around the world. One of the most notable features of Alzheimer's disease is that affected individuals can live for extended periods of time (ten or more years) while being in an extremely debilitated state often requiring round the clock care. Accordingly, the disease takes not only an enormous emotional toll, but also exacts a tremendous financial toll on affected individuals and their families. Therapies which improve neuronal function have substantial utility in improving the quality of life of Alzheimer's sufferers.

[0360] Huntington's disease is a degenerative disease whose symptoms are caused by the loss of cells in a part of the brain called the basal ganglia. This cell damage affects cognitive ability (thinking, judgment, memory), movement, and emotional control. Symptoms appear gradually, usually in midlife, between the ages of 30 and 50. However, the disease can also strike young children and the elderly. Huntington's disease is a genetic disorder. Although people diagnosed with the disease can often maintain their independence for several years following diagnosis, the disease is degenerative and eventually fatal. Currently, there are no treatments available to either cure or to ameliorate the symptoms of this disease. Furthermore, the onset of Huntington's disease is typically in middle-age (approx age 40), at a time when many people have already had children. Thus, people have usually passed this fatal genetic disorder to their off-spring before they realize that they are ill.

[0361] Amyotrophic lateral sclerosis (ALS), often referred to as "Lou Gehrig's disease," is a progressive neurodegenerative disease that attacks motor nerve cells in the brain and the spinal cord. Degeneration of motor neurons affect the ability of the brain to initiate and control muscle movement. With all voluntary muscle action affected, patients in the later stages of the disease become totally paralyzed, and eventually die.

[0362] Multiple sclerosis (MS) is an illness diagnosed in over 350,000 persons in the United States today. MS is characterized by the appearance of more than one (multiple) areas of inflammation and scarring of the myelin in the brain and spinal cord. Thus, a person with MS experiences varying degrees of neurological impairment depending on the location and extent of the scarring. The most common characteristics of MS include fatigue, weakness, spasticity, balance problems, bladder and bowel problems, numbness, vision loss, tremor and vertigo. The specific symptoms, as well as the severity of these symptoms, varies from patient to patient and is largely determined by the particular location within the brain of the lesions.

[0363] MS is considered an autoimmune disease. Recent data suggest that common viruses may play a role in the onset

of MS. If so, MS may be caused by a persistent viral infection or alternatively, by an immune process initiated by a transient viral infection in the central nervous system or elsewhere in the body. Epidemiological studies indicating the distribution of MS patients suggest that there is a triggering factor responsible for initiating onset of the disease. Without being bound by theory, it appears that some environmental factor, most likely infectious, must be encountered.

[0364] The incidence of MS is higher in North America and Europe and this geographic distribution is further suggestive of an environmental influence(s) underlying onset of MS. Additionally, MS is more prevalent in women than in men, and is more common amongst Caucasians than within either Hispanic or African-American populations. Interestingly, MS is extremely rare within Asian populations.

[0365] Macular degeneration is a catch-all term for a number of different disorders that have a common end result: the light-sensing cells of the central region of the retina—the macula—malfunction and eventually die, with gradual decline and loss of central vision, while peripheral vision is retained. Most cases of macular degeneration are isolated, individual, occurrences, mostly in people over age 60. These types are called Age Related Macular Degeneration (AMD). More rarely however, younger people, including infants and young children, develop macular degeneration, and they do so in clusters within families. These types of macular degeneration are collectively called Juvenile Macular Degeneration and include Stargardt's disease, Best's vitelliform macular dystrophy, Doyne's honeycomb retinal dystrophy, Sorsby's fundus dystrophy, Malattia levantine, Fundus flavimaculatus, and Autosomal dominant hemorrhagic macular dystrophy.

[0366] The present invention makes available effective therapeutic agents for restoring cartilage function to a connective tissue. Such methods are useful in, for example, the repair of defects or lesions in cartilage tissue which is the result of degenerative wear such as that which results in arthritis, as well as other mechanical derangements which may be caused by trauma to the tissue, such as a displacement of torn meniscus tissue, meniscectomy, a Taxation of a joint by a torn ligament, misalignment of joints, bone fracture, or by hereditary disease. The present reparative method is also useful for remodeling cartilage matrix, such as in plastic or reconstructive surgery, as well as periodontal surgery. The present method may also be applied to improving a previous reparative procedure, for example, following surgical repair of a meniscus, ligament, or cartilage. Furthermore, it may prevent the onset or exacerbation of degenerative disease if applied early enough after trauma.

[0367] Such connective tissues as articular cartilage, interarticular cartilage (menisci), costal cartilage (connecting the true ribs and the sternum), ligaments, and tendons are particularly amenable to treatment. As used herein, regenerative therapies include treatment of degenerative states which have progressed to the point of which impairment of the tissue is obviously manifest, as well as preventive treatments of tissue where degeneration is in its earliest stages or imminent. The subject method can further be used to prevent the spread of mineralisation into fibrotic tissue by maintaining a constant production of new cartilage.

[0368] In an illustrative embodiment, the subject method can be used to treat cartilage of a diarthroidal joint, such as a knee, an ankle, an elbow, a hip, a wrist, a knuckle of either a finger or toe, or a temporomandibular joint. The treatment can

be directed to the meniscus of the joint, to the articular cartilage of the joint, or both. To further illustrate, the subject method can be used to treat a degenerative disorder of a knee, such as which might be the result of traumatic injury (e.g., a sports injury or excessive wear) or osteoarthritis.

[0369] In still further embodiments, agents of the present invention can be employed for the generation of bone (osteogenesis) at a site in the animal where such skeletal tissue is deficient. For instance, administration of an agent that promotes the differentiation of stem cells to bone can be employed as part of a method for treating bone loss in a subject, e.g. to prevent and/or reverse osteoporosis and other osteopenic disorders, as well as to regulate bone growth and maturation. For example, preparations comprising the identified agents can be employed, for example, to induce endochondral ossification. Therapeutic compositions can be supplemented, if required, with other osteoinductive factors, such as bone growth factors (e.g. TGF- β factors, such as the bone morphogenetic factors BMP-2 and BMP-4, as well as activin), and may also include, or be administered in combination with, an inhibitor of bone resorption such as estrogen, bisphosphonate, sodium fluoride, calcitonin, or tamoxifen, or related compounds.

[0370] The present invention provides methods and compounds that can be used to promote regeneration, for example, regeneration of endodermally derived cells, tissues, and organs. Such methods and compositions can be used to treat conditions associated, in whole or in part, by loss of, injury to, or decrease in functional performance of endodermal cell types. By way of example, definitive endodermal cell type include, but are not limited to, hepatocytes of the liver, pancreatic cell types such as β -islet cells, cells of the lung, and cells of the gastrointestinal tract. The following are illustrative of disease states that can be treated using agents that promote regeneration of specific endodermal cell types.

Pancreatic Diseases

1. Diabetes Mellitus

[0371] Diabetes mellitus is the name given to a group of conditions affecting about 17 million people in the United States. The conditions are linked by their inability to create and/or utilize insulin. Insulin is a hormone produced by the beta cells in the pancreas. It regulates the transportation of glucose into most of the body's cells, and works with glucagon, another pancreatic hormone, to maintain blood glucose levels within a narrow range. Most tissues in the body rely on glucose for energy production.

[0372] Diabetes disrupts the normal balance between insulin and glucose. Usually after a meal, carbohydrates are broken down into glucose and other simple sugars. This causes blood glucose levels to rise and stimulates the pancreas to release insulin into the bloodstream. Insulin allows glucose into the cells and directs excess glucose into storage, either as glycogen in the liver or as triglycerides in adipose (fat) cells. If there is insufficient or ineffective insulin, glucose levels remain high in the bloodstream. This can cause both acute and chronic problems depending on the severity of the insulin deficiency. Acutely, it can upset the body's electrolyte balance, cause dehydration as glucose is flushed out of the body with excess urination and, if unchecked, eventually lead to renal failure, loss of consciousness, and death. Over time, chronically high glucose levels can damage blood vessels, nerves, and organs throughout the body. This can lead to other

serious conditions including hypertension, cardiovascular disease, circulatory problems, and neuropathy.

2. Pancreatitis

[0373] Pancreatitis can be an acute or chronic inflammation of the pancreas. Acute attacks often are characterized by severe abdominal pain that radiates from the upper stomach through to the back and can cause effects ranging from mild pancreas swelling to life-threatening organ failure. Chronic pancreatitis is a progressive condition that may involve a series of acute attacks, causing intermittent or constant pain as it permanently damages the pancreas.

[0374] Normally, the pancreatic digestive enzymes are created and carried into the duodenum (first part of the small intestine) in an inactive form. It is thought that during pancreatitis attacks, these enzymes are prevented or inhibited from reaching the duodenum, become activated while still in the pancreas, and begin to autodigest and destroy the pancreas. While the exact mechanisms of pancreatitis are not well understood, it is more frequent in men than in women and is known to be linked to and aggravated by alcoholism and gall bladder disease (gallstones that block the bile duct where it runs through the head of the pancreas and meets the pancreatic duct, just as it joins the duodenum). These two conditions are responsible for about 80% of acute pancreatitis attacks and figure prominently in chronic pancreatitis. Approximately 10% of cases of acute pancreatitis are due to idiopathic (unknown) causes. The remaining 10% of cases are due to any of the following: drugs such as valproic acid and estrogen; viral infections such as mumps, Epstein-Barr, and hepatitis A or B; hypertriglyceridemia, hyperparathyroidism, or hypercalcemia; cystic fibrosis or Reye's syndrome; pancreatic cancer; surgery in the pancreas area (such as bile duct surgery); or trauma.

Acute Pancreatitis

[0375] About 75% of acute pancreatitis attacks are considered mild, although they may cause the patient severe abdominal pain, nausea, vomiting, weakness, and jaundice. These attacks cause local inflammation, swelling, and hemorrhage that usually resolves itself with appropriate treatment and does little or no permanent damage. About 25% of the time, complications develop, such as tissue necrosis, infection, hypotension (low blood pressure), difficulty breathing, shock, and kidney or liver failure.

Chronic Pancreatitis

[0376] Patients with chronic pancreatitis may have recurring attacks with symptoms similar to those of acute pancreatitis. The attacks increase in frequency as the condition progresses. Over time, the pancreas tissue becomes increasingly scarred and the cells that produce digestive enzymes are destroyed, causing pancreatic insufficiency (inability to produce enzymes and digest fats and proteins), weight loss, malnutrition, ascites, pancreatic pseudocysts (fluid pools and destroyed tissue that can become infected), and fatty stools. As the cells that produce insulin and glucagons are destroyed, the patient may become permanently diabetic.

3. Pancreatic Insufficiency

[0377] Pancreatic insufficiency is the inability of the pancreas to produce and/or transport enough digestive enzymes to break down food in the intestine and allow its absorption. It

typically occurs as a result of chronic pancreatic damage caused by any of a number of conditions. It is most frequently associated with cystic fibrosis in children and with chronic pancreatitis in adults; it is less frequently but sometimes associated with pancreatic cancer.

[0378] Pancreatic insufficiency usually presents with symptoms of malabsorption, malnutrition, vitamin deficiencies, and weight loss (or inability to gain weight in children) and is often associated with steatorrhea (loose, fatty, foul-smelling stools). Diabetes also may be present in adults with pancreatic insufficiency.

Liver Diseases

1. Hepatitis

[0379] There are two major forms of hepatitis: one in which the liver is damaged quickly (called acute hepatitis) and one in which the liver is damaged slowly, over a long time (called chronic hepatitis). Hepatitis can be caused by chemicals, however, it is most commonly due to infection by one of several viruses that mainly damage the liver, termed hepatitis viruses. These viruses have been named in the order of their discovery as hepatitis A, B, C, D, and E. Hepatitis A is spread through infected water and food and is especially common in children. Most infected people don't even know they have been exposed to the virus. Hepatitis B is fairly common, especially in Asia and Africa. Although hepatitis B is less common in other parts of the world, it is still the most common cause of acute viral hepatitis in North America and Europe. Hepatitis B can be spread by exposure to blood, through sexual relations, and during pregnancy and childbirth. Symptoms of hepatitis B may be absent, mild and flu-like, or acute. Approximately 1-3% of patients become chronically infected, able to continue to infect others, and often have chronic damage to the liver. Those with weakened or compromised immune systems are at an increased risk to become carriers (about 10%). Newborns are especially vulnerable, with over 90% becoming carriers. Hepatitis C is passed the same way as hepatitis B. Hepatitis C is less common than B as a cause of acute hepatitis, but the majority of the people who contract it become chronically infected, able to spread the infection to others, and usually have chronic damage to the liver. Hepatitis D and E are rare in the United States, however, they are responsible for liver damage elsewhere in the world.

2. Cirrhosis

[0380] Anything that causes severe ongoing injury to the liver can lead to cirrhosis. It is marked by cell death and scar formation and is a progressive disease that creates irreversible damage. Cirrhosis has no signs or symptoms in its early stages, but as it progresses, it can cause fluid build-up in the abdomen (called ascites), muscle wasting, bleeding from the intestines, easy bruising, enlargement of the breasts in men (called gynecomastia), and a number of other problems.

3. Obstruction

[0381] Gallstones, tumors, trauma, and inflammation can cause blockage or obstructions in the ducts draining the liver (bile ducts). When an obstruction occurs, bile and its associated wastes accumulate in the liver and the patient's skin and eyes often turn yellow (jaundice). Bilirubin accumulating in

the urine turns it a dark brown color, while lack of bilirubin in the intestines causes the stool to become very pale colored.

[0382] Obstruction of the hepatic vein, the vein from the liver, may also occur, reducing blood flow out of the liver. This obstruction may be due to tumors pushing against the vein or from blood clot formation within the vein. Obstructions may be chronic and cause few symptoms, but they can also be acute and life threatening. Some can be treated with medications; others require surgery.

4. Fatty Liver

[0383] Fatty liver causes liver enlargement, tenderness, and abnormal liver function. The most common cause is excessive alcohol consumption. Another cause of fatty liver is NASH (nonalcoholic steatohepatitis). While symptom of fatty liver are often fairly mild, the condition can lead to chronic hepatitis and cirrhosis.

5. Genetic Liver Disorders

[0384] Hemochromatosis is the most common genetic liver disorder. It involves excess iron storage and is usually diagnosed in adults. There are numerous genetic liver diseases that affect children. Most of the diseases involve a defective element that results in liver injury (such as biliary atresia, where the bile ducts are absent or too small) or a missing enzyme or protein that leads to damaging deposits in the liver (such as galactosemia, the absence of a milk sugar enzyme, which leads to milk sugar accumulation; and Wilson's disease, where copper builds up in the liver).

[0385] Liver disease is often discovered during routine testing. It may not cause any symptoms at first or the symptoms may be vague, like weakness and loss of energy. In acute liver disease, symptoms related to problems handling bilirubin, including jaundice (yellowing of the skin and eyes), dark urine, and light stools, along with loss of appetite, nausea, vomiting, and diarrhea are the most common. Chronic liver disease symptoms include jaundice, dark urine, abdominal swelling (due to ascites), pruritus (itching), unexplained weight loss or gain, and abdominal pain.

(v) Detection Methods

[0386] Direct biophysical measurements can be obtained in any of a number of ways. Electrophysiological techniques are the classic approach used to measure bioelectrical phenomena, often using KCl-filled microelectrodes connected via Ag—AgCl junctions to a very high input impedance preamplifier (to avoid draining current from the system) to measure voltage levels within cells or beneath epithelia. Intracellular recording and voltage clamping are used to measure membrane voltage and whole cell currents. Patch recording and patch clamping can be used to measure ion flux through a limited number of channels in a small area of membrane. These techniques are accepted and powerful, and have been reviewed extensively (Jurkat-Rott and Lehmann-Horn, 2004, Curr Pharm Biotechnol 5: 387-395; Park et al., 2002, Pflugers Arch 444: 305-316). A newer tool, the self-referencing “vibrating” ion-selective probe (SERIS) is a non-invasive technique for detecting and measuring ion gradients at the surface of cells (Hotary et al., 1992, Development 114: 985-996; Nuccitelli, 1980, Federation Proceedings 39: 2129; Nuccitelli, 1987, Biophysical Journal 51: A447; Smith et al., 1999, Microscopy Res & Techniques 46: 398-417; Smith and Trimarchi, 2001, Am J of Physiology-Cell Physiology 280:

C1-C11). The technique utilizes ion-specific ionophore-filled microelectrodes. The tip of the electrode is vibrated at about 300 Hz between two points about 10 μm apart, one closer one farther away from a cell's plasma membrane. A difference in concentration at the two points indicates a gradient, which is measured quantitatively. An ion gradient near the outer surface of a cell implies transport of that ion across that membrane (without transport the gradient would quickly dissipate). This technique is powerful and allows the characterization of ion flux that can be used to infer the physiological state of the cell, as well as reveal the bioelectrical signals that the given cell or tissue is sending to its neighbors.

[0387] Fluorescent and other detectable ion-reporting can also be used (Amirand et al., 2000, Biol Cell 92: 409-419; Bassnett, 1990, Am J Physiol 258: C171-178; Bassnett, 1990 J Physiol 431: 445-464; Dascalu et al., 1993, J Physiol 461: 583-599; De Clerck et al., 1994, J Immunol Methods 172: 115-124; Edwards et al., 1998, Hum Reprod 13: 3441-3448; Epps et al., 1994, Chem Phys Lipids 69: 137-150; Franck et al., 1996, J Biotechnol 46: 187-195; Gasalla-Herraiz et al., 1995, Biochem Biophys Res Comm 214: 373-388; Krotz et al., 2004, Arterioscler Thromb Vasc Biol 24: 595-600; Sater et al., 1994, Development 120: 433-442). Such dyes are also referred to throughout the present application as 'voltage sensitive agents', 'voltage sensitive agents that produce a detectable signal', and 'pH sensitive agents.' Ion-reporting agents are readily available. They can be used to monitor (i) ion flux of virtually any ion species, (ii) membrane voltage, and (iii) intracellular pH. The ion-sensing ability of these dyes are based on the principle that binding of an ion changes the conformation of the molecule sufficiently to alter its fluorescence spectrum, a phenomenon referred to as a "spectral shift" or a "spectral response". This is illustrated in FIG. 8. Each of the curves illustrated in FIGS. 8A-8F is a spectrum of the same (imaginary) dye. The curve is made by measuring the intensity of the light emitted by the dye at two different wavelengths. As the concentration of the sensed ion ("[Ion]") changes, the curve moves, i.e. the spectrum changes. This change in the spectrum is known as the spectral shift. A graph of the spectral shift can be used as a standard curve.

[0388] Voltage sensitive agents can be loaded into cells by injection, electroporation, or, most commonly, by soaking the cells in membrane permeant (acetoxyethyl—AM) forms of the agent. When acetoxyethyl forms of voltage sensitive agents are internalized by cells, esterases present intracellularly cleave off the AM moiety, thus trapping the active form of the dye inside cells.

[0389] A voltage sensitive agent dye can be used for ion sensing if the bound and unbound states fluoresce differently. Referring to FIG. 8A, note that when excited by 488 nm light, this imaginary dye fluoresces most intensely at about 555 nm in the presence of 10^{-8} M ion (when fewer dye molecules are bound), whereas it fluoresces most intensely at about 585 nm in the presence of 10^{-6} M ion (when more dye molecules are bound). The units of emission intensity are not important. The scale can be calibrated depending on whether the fluorescence is evaluated using a confocal microscope, a dissecting microscope fitted with fluorescent filters, a fluorimeter, or a cytometer. Regardless of the particular scale or method used to evaluate the fluorescence, the important point is there is a measurable difference between the curves.

[0390] There are two ways that a fluorophore can "fluoresce differently": it's emission, when excited at a single wave-

length, can be ion-concentration sensitive (FIG. 8A) or its wavelength of maximal excitation, when its emission is monitored at a single wavelength, can be ion-concentration sensitive (FIG. 8D). cSNARF-1, an H⁺-reporting dye used for monitoring pH and/or H⁺ flux, is a so-called "dual emission" dye. Fluo-3, a widely used Ca²⁺-reporting dye, is a "dual excitation" dye. The particular agent selected depends on the model system, as well as the equipment available to analyze the your results.

[0391] In certain embodiments, the ion-reporting dyes are ratiometric. The ratiometric dyes allow the user to correct for artifacts that may affect the spectral shift, such as bleaching, differences in dye concentration or cell thickness, and spatial variation in instrument sensitivity. The easiest way to correct for local conditions is to measure them, and divide out their influence. That is, take a ratio: the intensity of the wavelength of interest (called λ_1) over the intensity of the local-condition light. To get the latter, you measure the same spot (the same dye molecules, in fact) at a second wavelength (called λ_2). If one is using a dual emission dye, emitted light of a different wavelength is collected. If one is using a dual excitation dye, light emitted at the same wavelength, but excited by a different wavelength is collected.

[0392] For example, using the imaginary dual-emission dye of FIG. 8A, one could illuminate with 488 nm light, then collect 555 nm light and 598 nm light (FIG. 8B). By way of further example using the imaginary dual-excitation dye of FIG. 8D, one would excite with 480 nm light and collect 545 nm light, then excite with 514 nm light and collect 545 nm light (FIG. 8E). Using either approach, producing two values of intensity: one for the wavelength that varies as a function of ion concentration (λ_1), and another that represents local conditions only (λ_2). In FIG. 8C, the intensity of the light emitted at \approx 555 nm (λ_1) is very sensitive to ion concentration, while the light emitted at \approx 590 (λ_2) is actually pH insensitive (this is called the isobestic point of the dye). In FIG. 8E, the intensity of the 545 nm light emitted when the dye is excited at 514 nm is very sensitive to ion concentration (λ_1), while at 480 nm, it is not (λ_2). The great advantage of using the ratio λ_1/λ_2 (often referred to as R) is that the two intensities will be similarly affected by local conditions, meaning their ratio will only be sensitive to ion concentration.

[0393] FIG. 8C represents the calibration (or standard curve) for the imaginary dual emission dye. The three points represent R from each of the three curves shown in 8B: 120/80=1.5, 160/80=2.0, 200/80=2.5. They are graphed as a function of the negative log of the concentration (pIon; if the ion were H⁺, this would be the pH scale). FIG. 8F is the calibration resulting from taking the ratios for the dual excitation dye, i.e. the data from FIG. 8E. The S-curves drawn through those points illustrate a critical point, which is that any given dye is only ion-concentration sensitive in a defined range of concentrations; above that concentration the dye is all bound, so the spectrum can not shift anymore, while below that concentration, the dye is all unbound, with the same result. The pK_d of the dye is the value of ion concentration that gives a ratio exactly half way between the completely bound and completely unbound states: for the imaginary dual-emission dye, pK_d is approximately 7.2; for the dual-excitation dye, pK_d is approximately 6.2. This value will be available from the manufacturer. The concentrations one wishes to measure must be near the pK_d of the dye so that the dye is ion-concentration sensitive.

[0394] Probably the most familiar of the ratiometric ion-sensing fluorescent probes are the calcium indicators, including the Fluo, Fura and INDO dyes. pH sensitive dyes, such as BCECF and the SNARF dyes, are also available. There are also probes for many other ions including Mg²⁺, Na⁺, K⁺, and Cl⁻. The Molecular Probes catalog is an invaluable resource for exemplary reagents. The invention contemplates the use of agents including, but not limited to, ion sensing dyes that can be used to indicate ion flux of a particular species of ion. The invention further contemplates the use of voltage sensitive agents (e.g., voltage sensitive reporting dyes) and pH sensitive agents (e.g., pH sensitive dyes).

[0395] In addition to ion sensing dyes, there are also V_m reporting dyes. These fluorophores undergo a spectral shift in response to V_m. There are two categories of V_m dyes, those that report fast changes, such as action potentials, and those that react more slowly, reporting V_m averages over longer periods of time. The fast-response dyes, such as Di-8-ANEPPS, localize to the membrane, and undergo a spectral shift due to a redistribution of intramolecular charge caused by a change in V_m. Thus, their spectral shift occurs quickly and can be used as a measure of fast changes in V_m. The slow-response dyes, in contrast, are anionic or cationic molecules that accumulate inside the cell due, it is thought, to an electrophoretic mechanism driven by the voltage across the membrane. The spectrum of these dyes is affected by both environment (intra- vs. extracellular) and concentration. For a cationic dye, such as TMRE, persistent hyperpolarization will cause the accumulation of more dye molecules, and the intensity of the shifted light will increase. For an anionic dye, such as DiBAC₄(3), persistent depolarization will have the same effect. Therefore, one consideration when choosing a dye is the expected voltage, and the expected direction of its change. In certain embodiment, experiments can be conducted in parallel using both an anionic and a cationic dye. In such embodiments, detecting opposite changes in the respective intensities of the two dyes can serve as a good control. Because the spectral shift of the slow-response probes depends on the movement of molecules, these dyes are appropriate for measuring longer-term phenomena, such as changes in the resting potential.

[0396] In one embodiment of any of the foregoing, the voltage sensitive agent produces a detectable fluorescent signal. Exemplary voltage sensitive agents include, but are not limited to, bis-(1,3-dibutylbarbituric acid)pentamethine oxonol (DiBAC₄(5)); bis-(1,3-dibutylbarbituric acid)trimethine oxonol (DiBAC₄(3)); bis-(1,3-diethylthiobarbituric acid)trimethine oxonol (DiSBAC₂(3)); 3,3'-diethyloxacarbocyanine iodide (DiOC₂(3)); 3,3'-diheptyloxacarbocyanine iodide (DiOC₇(3)); 3,3'-dihexyloxacarbocyanine iodide (DiOC₆(3)); 3,3'-dipentyloxacarbocyanine iodide (DiOC₅(3)); 1,1',3,3',3'-hexamethylindodicarbocyanine iodide (DiIC₁(5)); a structural variant thereof; or a functional variant thereof.

[0397] In another embodiment, the pH sensitive agent produces a detectable fluorescent signal. Exemplary pH sensitive agents include, but are not limited to, 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein (BCECF); 5-(and-6)-carboxy SNARF®-1; Lysotracker® Blue DND-22; Lysotracker® Green DND-26; Lysotracker® Red DND-99; Lysotracker® Yellow HCK-123; a structural variant thereof; or a functional variant thereof.

[0398] In another embodiment, the agent is an ion sensing dye. Ion sensing dyes capable of detecting ion flux of species including, but not limited to, Ca²⁺, H⁺, Cl⁻, K⁺, Na⁺, and Mg²⁺ can be used.

[0399] In addition to the foregoing methods, some particular reagents exist for further analyzing the role of gap junctions in a particular biological process. Junctional paths can be traced by microinjection of fluorescent small molecule dyes in large cells (Guthrie et al., 1988, Development 103: 769-783) or scrape-loaded into smaller cells. Various fluorescent techniques such as FRAP and photo-bleaching/uncaging of permeable molecules can be used to study GJC in vivo (Bedner et al., 2003, Exp Cell Res 291: 25-35; Braet et al., 2003, Cell Calcium 33: 37-48; Lee et al., 1995, Glia 15: 195-202; Pappas et al., 1996, Glia 16: 7-15; Suadicani et al., 2004, Glia 48: 217-229). A system of two differentially fluorescent molecules can be used to evaluate whether gap junctional communication between two cells exists. For example, two different dyes—a larger MW dye which does not pass through gap junctions (such as Molecular Probes' "Rhodamine-linked 10 kDa Dextran") together with a small (<1 kD) tracer—can be used. The inclusion of the larger dye controls for artifacts since only cell pairs which show transfer of the small dye but not the large dye represent a true instance of GJC. The permeability of the specific gap junctions involved in your system can be tested using a panel of fluorescent small molecule probes with different shape/charge/size characteristics. The following exemplary agents can be used: Lucifer yellow (MW=443, charge of -2); 2',7'-dichlorofluorescein (MW=401, charge of -1); neurobiotin (MW=287, charge of +1); 6-carboxyfluorescein (MW=376, charge of -2); DAPI (MW=350, charge of +1); ethidium bromide (MW=314, charge of +1); propidium iodide (MW=414, charge of +2); biocytin (MW=373, charge of 0); biotin-X cadaverin (MW=442, charge of +1); alexa 350 hydrazide (MW=349, charge of -1); alexa 488 hydrazide (MW=570, charge of -1). These and other agents are described in Meda, 2000, Methods 20: 232-244.

EXEMPLIFICATION

[0400] The invention now being generally described, it will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention and are not intended to be limiting in any way.

Example 1

Identification of a Role for Ion Flux in a Biological Process

Left-Right Asymmetry

[0401] We performed a compound screen to identify an ion transporter protein or a class of ion transporter proteins involved in left-right asymmetry. We employed a candidate approach to identify a manageable number of promising candidates for further molecular analysis. Briefly, groups of *Xenopus* embryos were treated from fertilization to stage 7 with inhibitors of various pumps, channels, and other ion transporters. Embryos cultured in the presence of compound were assayed to assess whether the compound had an effect on left-right asymmetry. Specifically, reversals of the heart, gut, or gall-bladder were assessed by morphological inspection at stage 45.

[0402] All of the reagents used in this screen were selected on the basis of high specificity for known electrogenic targets, and were titrated to ensure that the DAI (dorsoanterior index) of the treated embryos was normal, thus avoiding confounding randomization caused by midline defects. Of the compounds evaluated, various inhibitors of the H⁺-V-ATPase all induced strong independent randomization of the sidedness of the 3 organs assayed (e.g., compounds that specifically inhibited the activity of this ion transporter protein induced heterotaxia).

[0403] For example, the potent and highly-specific V-ATPase blockers concanamycin and bafilomycin induced heterotaxia in 35% (n=100) and 36% (n=100) of the embryos, respectively. In all positive cases, embryos were free of generalized toxicity or gastrulation defects. Importantly, inhibition of the other two major classes of H⁺ transporters, carbonic anhydrase and the sodium-hydrogen exchanger (NHE), did not affect laterality. Taken together, specific randomization of asymmetry, in the absence of disturbances in anterior-posterior or dorsal-ventral patterning, was induced by inhibitory compounds targeting the H⁺-V-ATPase, but not by compounds that modulated the activity of other H⁺ pumps or of a wide range of other ion transporters. These data implicate endogenous H⁺-V-ATPase in this particular biological process in *Xenopus* embryos.

[0404] We confirmed the screen results by characterizing the function of the H⁺-V-ATPase using molecular reagents. To unequivocally test the requirement for H⁺-V-ATPase activity, we analyzed left-right asymmetry in embryos injected with YCHE78 mRNA (e.g., we overexpressed this nucleic acid into embryos). This nucleic acid encodes a well-characterized dominant negative H⁺-V-ATPase subunit E (Lu et al., 2002, J Biol Chem 277: 38409-38415). Misexpression of this dominant negative construct specifically induced 20% heterotaxia ([n=191], controls=4% [n=422], $\chi^2=40.6$, p<<0.001). In contrast, injection of mRNAs encoding unrelated ion transporter proteins—the *Xenopus* H,K-ATPase α subunit or the dominant negative Kir2.2 subunit—did not cause heterotaxia (1% heterotaxia each, n=85 and 93, respectively). These data are consistent with the screen results, and support a role for ion flux mediated by a specific ion transporter protein—H⁺-V-ATPase—in LR patterning.

[0405] As outlined above, effects on asymmetry were initially evaluated by inspection of embryo morphology. To extend these results, we examined expression of transcripts known to be involved in establishing the embryonic left-right axis during development. Embryos were treated with the H⁺-V-ATPase inhibitory compound concanamycin. In addition to morphologically observable effects on asymmetry, perturbation of H⁺ flux by inhibiting the function of H⁺-V-ATPase perturbed the normally left-sided gene expression of the molecular marker XNr-1.

[0406] To further characterize a role for this specific ion transporter protein (H⁺-V-ATPase), we examined its expression. We performed immunohistochemical analysis of early embryo sections to determine whether H⁺-V-ATPase localization is consistent with a role in asymmetry. We first characterized an antibody successfully used to specifically detect subunit A of the H⁺-V-ATPase in mammalian tissues. Western blot analysis on *Xenopus* early embryo extracts revealed a single clean band of the predicted size, and negative controls exhibited no visible signal. Immunohistochemistry with this antibody revealed that at the 2-cell stage, H⁺-V-ATPase subunits exhibited “fingers” of localization extending up from a

pool in the vegetal cytoplasm into the animal half. At the four cell stage, staining is still asymmetrical, and is observed on the right side. The spatio-temporal pattern of expression of H⁺-V-ATPase subunits during early development is consistent with a role for this ion transporter protein in regulating early left-right asymmetry.

[0407] We additionally characterized the expression pattern of H⁺-V-ATPase subunits in chick and zebrafish embryos. In situ hybridization analysis of subunits A, C, and F, and immunocytochemistry for subunits F and C indicated that the H⁺-V-ATPase is expressed in the primitive streak of chick embryos at stages 2, 3 and 4, and in the node at stage 4. This spatio-temporal pattern of expression correlates with the time-period in which chick LR sidedness is determined (Levin et al., 2002, Cell 111: 77-89).

[0408] In zebrafish embryos, immunohistochemical analysis using antibodies to subunit C reveals expression in two- and four-cell stage embryos. Expression was also evaluated and detected later in development (e.g., the eight-cell stage, the thirty-two-cell stage, epiboly). Thus, H⁺-V-ATPase subunits are expressed at stages relevant to LR patterning in zebrafish.

Example 2

H⁺-V-ATPase-Dependent, Asymmetric H⁺ Flux

[0409] We examined ion flux using a self-referencing ion probe to measure H⁺ flux from living early blastomeres. The H⁺-V-ATPase endogenously acts to pump protons out of the cell (e.g., it mediates proton efflux). We note, however, that other ion transporters modulate influx of ions. The term ion flux refers to movement of ions across membranes, regardless of the direction. The techniques provided herein can be used to evaluate changes in either ion efflux or influx, and thus can generally be used to evaluate ion flux and changes in ion flux following manipulation of an ion transporter protein or a class of ion transporter proteins.

[0410] We detected a large net efflux of protons from the cleavage furrow of the two-cell stage (in *Xenopus* embryos). This efflux averaged 12.7 ± 22 pmole cm⁻² s⁻¹ (n=5) at about the midpoint of cleavage. Importantly, we also found evidence for asymmetry of H⁺-flux. As early as the four-cell stage, a distinct difference in proton efflux was detected across the ventral midline, with larger efflux occurring on the right side, consistent with the immunological localization of H⁺-V-ATPase subunits.

[0411] In measurements on 15 embryos made between the four-cell stage and stage 6, the average proton efflux from the middle of the right ventral quadrant was 4.1 ± 0.48 pmole cm⁻² s⁻¹ and from the left ventral quadrant was 1.9 ± 0.29 pmole cm⁻² s⁻¹. The average ratio of the efflux from the right side to efflux from the left side was 2.3 ± 0.3 .

[0412] To confirm that the asymmetric efflux was due to asymmetric H⁺-V-ATPase function, we compared H⁺ efflux in control embryos to efflux from embryos treated with a variety of ion flux inhibitors. Neither the absolute net proton fluxes nor their right/left ratios were affected by the application of Omeprazole (0.27 μ M), an inhibitor of H⁺/K⁺-ATPase (e.g., a different ion transporter protein). However, the highly specific H⁺-V-ATPase inhibitor concanamycin reduced the proton efflux on the right side to about half of its original value and eliminated the left/right asymmetry in the proton fluxes. Taken together, these data reveal the existence of consistent physiological asymmetry (H⁺ flux) in the four-cell

embryo, and confirm that the asymmetry in the flow of H⁺ ions out of the early blastomeres is due specifically to differential H⁺-V-ATPase activity.

[0413] Because in many systems H⁺-V-ATPase activity contributes significantly to the membrane voltage gradient V_{mem}, we also used the reporting dye DiBAC₄(3) to determine whether there is a consistent asymmetry to V_{mem}. Embryos marked with Alexa 647 dextran in the right ventral cell (RV) at the four cell stage (to allow orientation with respect to both axes) were monitored using DiBAC₄(3) at the 4, 8, 16, and 32 cell stages. At 16 cells, we observed hyperpolarization of the RV quadrant, relative to the LV quadrant. DiBAC₄(3) intensity on the left was greater than on the right. This pattern and its embryological timing are consistent with the self-referencing probe data, and with the hypothesis that higher net H⁺ efflux from the right side results in a difference in membrane potential across the ventral midline.

Example 3

Misexpression of an Ion Transporter Protein

[0414] To complement the experiments conducted using inhibitory compounds, we examined the effects of a gain-of-function treatment that would produce an excess, equal H⁺-flux across the plasma membrane on both sides of the midline. Because the H⁺-V-ATPase is a multi-subunit complex and may be difficult to re-constitute, we induced an ectopic H⁺ flux by expressing a well-characterized single-subunit plasma-membrane H⁺-pump, PMA1.2 (Masuda and Montero-Lomeli, 2000, Biochemistry and Cell Biol 78: 51-58). Use of this construct also allowed us to address whether it is the balance of H⁺ flux at the cell membrane that is important for LR asymmetry, since the H⁺-V-ATPase is also known to occur in vacuoles. In contrast, the PMA1.2 pump functions only in the cell membrane.

[0415] Misexpression of a H⁺ pump at cell surfaces throughout the embryo by microinjection of mRNA at the one-cell stage caused significant heterotaxia of embryos (PMA1.2=21% heterotaxic [n=135, unscored=22%, untreated=2% [n=187, unscored=3%], χ²=28.3, p<<0.001). These data are consistent with the importance of differential H⁺ flux in modulating left-right asymmetry.

Example 4

Manipulation of pH or Membrane Voltage Effects Asymmetry

[0416] We separately tested its two physiological roles—regulation of pH and membrane voltage gradients—by experimentally altering these two parameters independently of direct manipulations of the H⁺-V-ATPase. We first examined the effect of changing the pH of the embryo's external environment by raising or lowering the pH of the 0.1×MMR culture medium in which embryos were maintained. While neutral to high pH (7 to 11) had no effect on LR patterning, pH 5 to 6 caused a low level of LR patterning defects (6%, p=0.003), and pH 4 caused a significant level of heterotaxia (19%, p<<0.001). These data are consistent with the idea that a high external proton concentration will inhibit the activity of pumps, including H⁺-V-ATPase, that normally extrude H⁺ from the cytoplasm.

[0417] To alter embryonic pH without changing V_{mem}, we cultured embryos in the electroneutral OH⁻/C⁻ exchanger tributyltin chloride (TBT) to raise internal pH by increasing

OH⁻ influx. Treatment with 0.02 μM TBT from the one-cell stage to stage 13 caused heterotaxia in 17% of embryos ([n=151]; untreated 1% [n=361]; χ²=44.7, p<<0.001).

[0418] To study the effect of isolated changes in pH using greater molecular specificity, we overexpressed the electroneutral, plasma-membrane Na⁺/H⁺-exchanger NHE3 (Prætorius et al., 2000, American Journal of Physiology 278: G197-G206; Sabirov et al., 1999, J of Membrane Biol 172: 67-76). Injection of this antiporter mRNA within one hour of fertilization caused heterotaxia (NHE3=16% [n=77, unscored=32%], untreated=3% [n=114, unscored=5%], χ²=8.9, p=0.003) in the absence of non-specific toxicity. These findings confirm the importance of pH for LR asymmetry, and, similar to the PMA1.2 data, indicate that the relevant H⁺ flux occurs at the cell membrane, not in vesicles.

[0419] The H⁺-V-ATPase is electrogenic and can therefore significantly contribute to the steady-state V_{mem} in addition to affecting pH. We confirmed this in *Xenopus* blastomeres by imaging V_{mem} in vivo in control and concanamycin-treated embryos. Predicting that the membrane would depolarize if H⁺-V-ATPase-based proton efflux was inhibited, we monitored V_{mem} using the voltage-reporting dye DiBAC₄(3). As predicted, inhibition of H⁺-V-ATPase depolarizes V_{mem}.

[0420] To address the role of V_{mem} in the absence of pH changes, we altered V_{mem} directly by incubating embryos in the Na⁺/K⁺ inhibitor palytoxin (PTX), which converts the Na⁺/K⁺-ATPase into a non-specific ion channel thus dissipating the voltage gradient and depolarizing cells (Hilgemann, 2003, PNAS 100: 386-388). PTX treatment (2 nM) of embryos, from the one-cell stage to stage 6, caused abnormal LR patterning in 20% of treated embryos ([n=60, unscored=84%], untreated=2% [n=58, unscored=3%], χ²=8.3, p=0.004). We conclude that normal V_{mem}, or patterned differences in V_{mem}, are necessary for proper LR patterning.

Example 5

Examination of H⁺-V-ATPase Function in Other Model Systems

[0421] To determine whether chick H⁺-V-ATPase plays a role in LR asymmetry, we first examined the pattern of cellular pH in the early chick blastoderm. We then assessed the effect of inhibiting H⁺-V-ATPase activity using concanamycin. Using the pH reporting dye SNARF-1-AM (Buckler and Vaughan-Jones, 1990, Pfugers Archiv—European J of Physiology 417: 234-239; Morley et al., 1996, Biophysical Journal 70: 1294-1302), we found that area pellucida cells maintained higher pH than primitive streak cells in control embryos. As predicted by the expression of the H⁺-V-ATPase in the primitive streak, treatment with the H⁺-V-ATPase inhibitor concanamycin lowered the pH of cells in the streak.

[0422] We examined the effect of inhibiting H⁺-V-ATPase activity on the expression of the early left-sided marker Sonic hedgehog (Shh). Inhibition of H⁺-V-ATPase activity following treatment with concanamycin or DCCD specifically randomized the expression of Shh, as well as that of and the downstream marker Nodal. Compared to controls, in which these two markers are always left-sided, concanamycin induced aberrant sidedness of Shh expression in 67% and Nodal in 24% of embryos. Thus, as in *Xenopus*, H⁺-V-ATPase is upstream of the known asymmetric gene cascade, and is required for normal LR patterning in chicks. Perturbation of

this specific ion transporter protein using a compound that specifically inhibited ion flux mediated by this protein modulated left-right asymmetry.

[0423] We also examined the consequences of disrupting endogenous cell-membrane pH gradients in *Danio rerio* embryos. As reported above for *Xenopus*, microinjection into fertilized zebrafish eggs of mRNAs encoding YCHE78 (thus equalizing patterned H⁺ fluxes by loss-of-function mediated by the dominant negative H⁺-V-ATPase subunit E), the plasma membrane H⁺-pump PMA1.2, and the electroneutral Na⁺/H⁺ antiporter NHE3 (the latter constructs equalizing patterned H⁺ fluxes by ubiquitous over-expression of exogenous plasma membrane H⁺ pumps), all specifically randomized LR patterning of the visceral organs. Down-regulating H⁺-V-ATPase activity by injection of a dominant negative subunit E mRNA caused a high level of heterotaxic organ situs (YCHE78=36% [N=42, unscored=0%], untreated=5% [N=96, unscored=4%], $\chi^2=19.5$, p<<0.001). Injections of mRNA for NHE3, to alter cytoplasmic pH only, caused 34% heterotaxia, ([N=99, unscored=6%], untreated=5% [N=94, unscored=1%], $\chi^2=23.4$, p<<0.001). Expression of PMA1.2, to exogenously increase H⁺ flux, caused 29% heterotaxia, ([N=52, unscored=9%], untreated=5% [N=96, unscored=4%], $\chi^2=14.2$, p<<0.001).

[0424] Taken together, these data strongly support the hypothesis that specific levels and/or distributions of H⁺-V-ATPase activity are necessary for correct LR asymmetry in zebrafish embryos. Compounds that specifically modulated the activity of this ion transporter protein modulated left-right asymmetry during embryonic development.

[0425] The results summarized in Examples 1-5 indicate a role for ion flux in a particular biological process. Furthermore, these results indicate that one can use the methods of the invention to identify a particular ion transporter protein or class of ion transporter proteins that mediate ion flux and thereby modulate a particular biological process. Once a particular ion transporter protein or class of ion transporter proteins are identified, expression analysis and functional analysis can be used to further analyze the role of the ion transporter protein during the biological process. Expression and functional studies can be used to extend the initial findings into other model systems or stages of development.

[0426] In Examples 1-5, the biological process was left-right asymmetry during embryonic development. One of the implicated targets, the H⁺/K⁺-ATPase, has been previously characterized in *Xenopus* and chick embryos (Levin et al., 2002, Cell 111: 77-89). We now present evidence for a role for H⁺ flux generally, and the H⁺-V-ATPase specifically, during the establishment of left-right asymmetry during embryonic development.

[0427] The H⁺-V-ATPase complex, more commonly referred to as the V-ATPase, is found in the membranes of vacuoles and other intracellular vesicles where it acidifies the intravesicular environment, thus activating enzymatic or other vesicle-dependent processes (Inoue et al., 2003, J Bioenerg Biomembr 35: 291-299; Kawasaki-Nishi et al., 2003, FEBS Letter 545: 76-85; Nishi, 2002, Nature Reviews Molecular Cell Biology 3: 94-103). In many cell types, including osteoclasts, kidney collecting duct, and chick chorioallantoic membrane, the H⁺-V-ATPase is also present in the plasma membrane (Baron et al., 1985, J Cell Biol 101: 2210-2222; Brown et al., 1987, J Cell Biol 105: 1637-1648; Klein et al., 1997, J Membrane Biology 157: 117-126; Narbaitz et al., 1995, Journal of Anatomy 186: 245-252; Nishi,

2002, Nature Reviews Molecular Cell Biology 3: 94-103; Schweikl et al., 1989, Journal of Biol Chem 264: 11136-11142), where, by pumping protons out of the cell, it affects cytoplasmic pH and the pH of the immediate extracellular environment (Brown and Breton, 2000, J Exp Biol 203: 137-145; Kawasaki-Nishi et al., 2003, FEBS Letter 545: 76-85; Morsomme and Boutry, 2000, Biochimica et Biophysica Acta 1465: 1-16; Nishi, 2002, Nature Reviews Molecular Cell Biology 3: 94-103; Scarborough, 2000, Cell Mol Life Sci 57: 871-883). The H⁺-V-ATPase is also electrogenic, and can contribute to the potential of the cell membrane in which it resides (Harvey, 1992, Physiology of V-ATPases. In V-ATPases, vol. 172, ed. W. Harvey and N. Nelson; Slack and Warner, 1975, Journal of Physiology 248: 97-120; Wieczorek, 1999, Bioessays 21: 637-648).

[0428] The H⁺-V-ATPase comprises two domains, the V0 and the V1, which are analogous, and largely homologous, to the F0 and F1 domains of mitochondrial ATPases. For every ATP molecule that is hydrolyzed by the V1, two protons are pumped across the membrane through proteins of the V0 domain (Nishi, 2002, Nature Reviews Molecular Cell Biology 3: 94-103).

METHODS: The following methods were used in the experiments summarized in Examples 1-5

[0429] For experiments using *Xenopus laevis*, embryos were collected according to standard protocols in 0.1× Modified Marc's Ringers (MMR) pH 7.8 supplemented with 0.1% Gentamicin. Embryo developmental stage was determined according to standard criteria (Nieuwkoop and Faber). For experiments using chicken embryos, standard pathogen-free white leghorn chick embryos from Charles River Laboratories (SPAFAS) were used. The embryos were maintained at 38° C., and developmental stage was determined according to standard criteria (Hamburger and Hamilton). For experiments using zebrafish embryos, *Danio rerio* embryos were obtained using standard methods and maintained at 28.5° C. in fish-system water, containing 1 drop per gallon Methyl-blue. Organ situs in fish embryos was evaluated at 5-7 days post-fertilization (dpf).

Assaying Organ Situs

[0430] *Xenopus* embryos at stage 45 were analyzed for position (situs) of three organs: the heart, stomach, and gallbladder. Only embryos with normal dorsoanterior development (DAI=5) and clear left- or right-sided organs were scored. Embryos with ambiguous (unscoreable) situs were also counted, and while not included in statistics, were used to gauge embryo quality and treatment toxicity. An embryo was considered heterotaxic if one, two, or all three organs were abnormally positioned. Experimental organ situs percentages were compared to that of untreated controls using a χ^2 test with Pearson correction (assuring increased stringency for significance). The autofluorescence of gallbladder and pancreas was used to score visceral situs in zebrafish embryos. Tricaine-anaesthetized 5-6 day larvae were examined on a Zeiss Stemi SV11 dissecting microscope under 488/40 nm illumination, using a 510 nm barrier filter. An embryo was considered heterotaxic if either or both organs were on the side opposite normal.

Pharmacological Treatments

[0431] *Xenopus* control embryos were incubated from 60 minutes post-fertilization to stage 6-7 in 8 to 10 ml of 0.1×

MMR/pH 7.8. For the screen, experimental embryos were incubated in 0.1×MMR containing compound. At stage 6-7 embryos were transferred to 0.1×MMR and maintained as described until scoring at stage 45. For all data shown, normal midline development and DAI were observed.

[0432] To perform a pharmacological screen on chick embryos with minimal disturbance to normal morphogenesis, we optimized a chicken *in ovo* culture system. A small hole was made on the top of each egg (prior to incubation), and 5 ml of light albumin was removed. The experimental solution, consisting of pharmacological compound in chicken light albumin and Pannett-Compton (PC) solution at a ratio of 5:1, was placed into the egg. Eggs were securely wrapped with Scotch tape, incubated at 37.5° C. to the desired stages, then fixed for analysis of laterality markers by *in situ* hybridization.

[0433] *Danio rerio* embryos were incubated at 28.5° C. in 25 ml of unadulterated or lobatomide A16- or concanamycin-containing fish water, from the one- to two-cell stage to 50% epiboly. Lobatomide A16 was used at 1 µM; concanamycin (Sigma) was used at 250 nM.

Self-Referencing Ion-Selective (Seris) Probe

[0434] The self-referencing ion-selective (Seris) probe consists of an ion selective microelectrode that is translated between two points in the extracellular space near the plasma membrane of a cell. If a net flux of ion into or out of the cell is detected, the resultant concentration gradient will be detected as a voltage difference from which the ion flux through the membrane can be calculated. In these studies, micropipettes with tip diameters of 3 µm were backfilled with a short (1-2 mm) column of 100 mM NaCl buffered with 10 mM HEPES at pH 7.0. Pressure from a syringe was applied to force the saline to the tip of the silanized pipette, the tip was inserted into a proton ionophore cocktail (Fluka Hydrogen Ionophore I-Cocktail B), and a short (15-20 µm) column of cocktail allowed to flow into the tip, displacing the saline. The electrodes were calibrated by moving between solutions of embryo culture medium (see below) of pH 6.5 and 7.5. Potassium selective electrodes were made using a valinomycin-based cocktail (Fluka Potassium Ionophore 1-Cocktail B), using a 100 µm column; the filling solution was 100 mM KCl. Calibration was performed in culture medium with K⁺ concentrations between 0.1 mM and 1.0 mM. Electrodes always gave at least a 58 mV change for a 10× change in solution K⁺.

[0435] Seris Measurements

[0436] For Seris measurements of proton flux, healthy *Xenopus* embryos were cultured in 0.1×MMR at pH 7.0. Embryos were illuminated from below with dark-field optics and from above with transmitted light, and viewed on a video monitor located outside the Faraday cage enclosing the electrophysiological apparatus. H⁺ flux measurements were made to the left and right, equidistant from the ventral midline, near the animal-vegetal (AV) midline, slightly on the animal side, and approximately in the middle of the left-ventral quadrant and the right-ventral quadrant of the embryo. The displacement of the electrode was along a line that lay 45° from the normal to the embryo's surface. This was true for both the left-ventral and right-ventral measuring position. Using the analysis of Arif et al. (Arif et al., 1995, Plant Cell Environment 18: 1319-1324), we calculate a correction factor of 105 for the presence of 0.5 mM HEPES buffer and have

applied that factor to generate the absolute flux values provided. The ratio of the fluxes across the ventral midline is unaffected by this correction.

Membrane Voltage Sensitive Dye DiBAC₄(3)

[0437] Bis-(1,3-dibarbituric acid)-trimethine oxanol (DiBAC₄(3), Molecular Probes, Eugene, Oreg.) is a membrane-permeant, fluorescent molecule that accumulates inside a cell in proportion to the membrane voltage. Because it is anionic, the more depolarized a cell, the greater the accumulation of DiBAC₄(3) and the greater the intensity of intracellular, relative to extracellular, fluorescence. Stock DiBAC₄(3) (1 mg/ml in DMSO) was diluted 1:10 in distilled water, then 1:100 in 0.1×MMR, for a final concentration of 1.9 µM. *Xenopus* embryos were soaked in dye solution for at least 30 minutes, then imaged submerged in dye, using a Leica TCS SP2 Spectral Confocal Imaging System mounted on a Leica upright DM RXE microscope. The dye was excited with 488 nm light from an argon laser and a 20 nm band of emission wavelengths centered at 515 was collected.

In Situ Hybridization

[0438] Whole mount *in situ* hybridization (WISH) on *Xenopus* was performed according to standard methods. DNA used to transcribe *in situ* hybridization probes were XNr-1 (Lohr et al., 1998, Developmental Genetics 23: 194-202; Lowe et al., 1996, Nature 381: 158-161); cShh (Levin et al., 1995, Cell 82: 803-814); cNodal (Levin et al., 1995, Cell 82: 803-814).

[0439] WISH analysis of zebrafish embryos was performed according to standard methods. Briefly, embryos were collected at the 20-22 somite stage, fixed in 4% PFA in phosphate buffered solution containing 0.1% Tween 20, and dehydrated in methanol. Embryos were rehydrated, digested with 10 µg/ml Proteinase K, and refixed in 4% PFA. Prehybridization was performed at 65° C. for 3 hours in 50% formamide, 5×SSC, 0.1% Tween-20, 4.6 mM citric acid, 50 µg/ml Heparin, and 500 µg/ml tRNA. Digoxigenin-labeled anti-sense riboprobes (Roche Applied Science, Penzberg, Germany) were added directly to pre-hybridization mix and allowed to incubate overnight at 65° C. Riboprobe was synthesized from linearized plasmid containing spaw cDNA. Washes were performed the following day at 65° C. in graded solutions from 100% hybridization mix to 100% 2×SSC, and then in 0.2× SSC. Embryos and α-dig antibody were pre-blocked at room temperature for at least 3 hours in a solution containing 1 part 10% Boehringer blocking reagent dissolved in 1M Maleic Acid, 1 part lamb serum, and 3 parts filtered Maleic Acid Buffer (MAB: 100 mM Maleic Acid, 150 mM NaCl, 0.1% Tween 20, 7.9 g/L NaOH, pH to 7.5). After pre-blocking, embryos were transferred to the α-dig antibody solution and blocked overnight at 4° C. The following morning embryos were washed first in MAB, and then in AP buffer solution (60 mM Tris HCL pH to 9.5, 60 mM NaCl, 30 mM MgCl₂, and 0.1% Tween-20). Staining was achieved with NBT and BCIP in AP buffer. After staining, embryos were refixed in 4% PFA, dehydrated in 100% methanol overnight to remove background staining, and stored in 80% Glycerol.

Immunocytochemistry

[0440] Immunocytochemistry was performed as described according to standard methods. Briefly, *Xenopus* embryos were fixed overnight at 4° C. in MEMFA; chick embryos were

fixed overnight in 4% PFA. Some embryos were embedded for sectioning at 40 μ m on a Leica Vibratome and others were left intact for whole mount analysis. After washing 3 \times in 1 \times PBS+0.1% Triton X-100 (PBST), samples were blocked with 20% Goat serum+0.2% BSA and incubated overnight at 4° C. with primary antibody at the following dilutions: anti-subunit A, 1:500; anti-myosin V, 1:500 (Chemicon #AB5887), anti-RFX3, 1:500; anti-subunit F, 1:500; anti-subunit c, 1:500 (antibody against peptide DAGVRGTAQ). After washing 6 \times with PBST, samples were incubated overnight at 4° C. with alkaline phosphatase conjugated secondary antibodies. After 6 more washings in PBST, detection was carried out using NBT and BCIP; the chromogenic reaction was stopped when the signal to noise ratio appeared optimal. Patterns reported for localization of V-ATPase subunits represent a consensus of data obtained from at least 15 embryos.

Microinjection of mRNA

[0441] *Xenopus* embryo mRNA injections were performed with capped, synthetic mRNA, according to standard protocols. Zebrafish embryo microinjection was performed using standard protocols. Three different constructs were used: YCHE78, PMA1.2, and NHE3. YCHE78 encodes a partial H⁺-V-ATPase subunit E (Lu et al., 2002, Journal of Biol Chem 277: 38409-38415) that acts as a dominant negative. PMA1.2 encodes a P-type H⁺-ATPase from *Saccharomyces cerevisiae* (Masuda and Montero-Lomeli, 2000, Biochemistry and Cell Biology 78: 51-58) which mimics the activity of H⁺-V-ATPase. NHE3 is an electroneutral proton pump (Soleimani et al., 1994, Biochim Biophys Acta 1195: 89-95; Couillon and Pouyssegur, 2000, J Biol Chem 275: 1-4). Results of injections are reported as: % of otherwise normal embryos that were heterotoxic; the sample size (n); when possible, the percent of injected embryos that died or were abnormal after gastrulation and thus not scored; the value of χ^2 and the p value comparing treated to controls.

Example 6

A Role for Ion Transporter Proteins in Dedifferentiation and Regeneration

[0442] The *Xenopus* embryo, a highly tractable vertebrate model system, regenerates its tail. The tail is thus an ideal model system for evaluating the effect of ion flux and membrane potential on regeneration. Within hours of amputation, the tail forms a blastema that then rapidly regenerates a perfect duplicate of the original tail. We used this system to evaluate the role of ion transporter proteins in regeneration. Furthermore, studies performed in the *Xenopus* tail are broadly applicable to studies of regeneration in multiple cell types because the tail is a complex appendage including somites (mesodermal structures), blood vessels, and neural tube. Other *Xenopus* structures that have an enhanced regenerative capacity include the limbs.

[0443] Having identified the *Xenopus* tail as an experimentally-tractable model of regeneration, we carried out several experiments to test whether bioelectric phenomena were relevant to this process. As previously reported by others, in the absence of any compounds or other manipulations, tails amputated at the midpoint between the posterior end of the hindgut and the caudal tip of the tail at approximately stage 40 regenerate with normal morphology. The process is robust (occurs normally in >90% of the embryos) and rapid. By stage 46, the tails of cut embryos are the same length (and shape) as uncut sibling controls (4 days at 22° C.).

[0444] To test whether ion flux was relevant to tail regeneration, we conducted a pharmacological screen using compounds that inhibit the activity of different types of ion channels, pumps, and co-transporters. We quantified the effects using a “regeneration index” which scores several specific aspects of regeneration and thus provides a continuous measure of regeneration.

[0445] In a first round of screening, we cultured embryos in medium containing a compound beginning immediately after tail amputation (stage 40-41). The following compounds had no effect on the rate or extent of tail regeneration: SCH28080 (inhibitor of H⁺/K⁺-ATPase); 9-anthracene-carboxylic acid (inhibitor of Cl⁻ channels); heptanol (blocker of gap junctional communications); diazoxide (opener of K⁺ channels); glibenclamide (inhibitor of K⁺ channels); ouabain (inhibitor of Na⁺/K⁺-ATPase); EIPA (inhibitor of Na⁺/H⁺ antiporter); THB (inhibitor of voltage-sensitive K⁺ channels); WAY-123 (blocker of ERG-K channels); amiloride (inhibitor of Na/H antiporter); 18- β -glycyrrhetic acid (blocker of gap junctional communication); EM12 (inducer of gap junctional communication). Additional screening identified further compounds that had no effect on the rate of tail regeneration (Table 2). Given that compounds that modulate ion flux mediated by the foregoing classes of ion transporters had no effect on the rate or extent of regeneration, we are not pursuing these classes of transporters in these biological processes.

[0446] FIG. 9 summarizes experiments performed using compounds that altered the regeneration index, and thus implicated a role for three classes of ion transporter proteins: K⁺ channels, the V-ATPase, and voltage-gated sodium channels. In contrast to amputated embryos exposed to vehicle, which regenerated normally, embryos exposed to blockers of the V-ATPase (e.g., concanamycin), K⁺ channels (BaCl), or voltage-gated Na⁺ (NaV) channels (tricaine) continued normal development in the absence of regeneration (FIG. 10).

[0447] These effects were specific to regeneration because (a) exposed embryos were alive and exhibited fairly typical developmental morphology, even in the absence of normal tail regeneration; and (b) despite the lack of tail regeneration, the wound created when the tail was amputated healed successfully.

[0448] These results demonstrated that tail regeneration depends on the activity of the three implicated target classes. The effect of BaCl strongly suggests a K⁺ channel. However, the lack of effect on regeneration observed in the presence of THB and glibenclamide suggests that the subject K⁺ channel is not a member of either the class of 6 transmembrane voltage-gated potassium channels or the class of inwardly-rectifying channels. The profile obtained from the compound screen suggests that K⁺ flux during tail regeneration is mediated by the class of two-pore channels (4 or 8 TM domains) of the KCNK family (Goldstein et al., 2001).

[0449] These results indicated that members of all three classes of ion transporter proteins are important during tail regeneration. Inhibition of any one of the three classes alone abolishes regeneration. However, the effect was not additive.

Example 7

A Role for Ion Transporter Proteins in Dedifferentiation and Regeneration

[0450] We next examined membrane voltage and pH during tail regeneration. We used the voltage-sensitive fluorescent dye DIBAC₄(3), and examined cultured embryos using

confocal microscopy (FIG. 11). Within 12 hours of amputation, the blastema cells exhibit a clear depolarization. Moreover, while the depolarization could be easily observed in blastema cells in embryos cultured without inhibitory compounds (FIG. 11C left), embryos in which regeneration was blocked by an inhibitory compound did not exhibit the same region of depolarization (FIG. 11C right). These observations support a model whereby amputation triggers the formation of a zone of depolarization which is necessary for subsequent regeneration steps.

[0451] We also detected a pH gradient at the *Xenopus* blastema. Based on our inhibitor studies, this pH gradient is likely driven by the V-ATPase (FIG. 11F).

Example 8

Expression Analysis of Ion Transporter Proteins During Tail Regeneration

[0452] We conducted expression analysis to evaluate the spatio-temporal expression of targets implicated in regeneration based on the compound screen. We performed immunohistochemical analysis using the following antibodies: an antibody raised against the V-ATPase subunit c'; a commercial NaV antibody (Sigma) and a KCNK1 antibody. We detected protein expression using each of these antibodies in *Xenopus* tail regenerates. Additionally, we evaluated proliferation, and observed a zone of proliferating cells in the blastema.

[0453] These results implicate three classes of ion transporter proteins in *Xenopus* tail regeneration: the V-ATPase H⁺ pump, a K⁺ channel, and a voltage-gated Na⁺ channel. When these targets are specifically inhibited after amputation, wound healing and normal embryonic development continues, while tail regeneration is absent. Moreover, using a voltage-sensitive fluorescent dye, we directly observed the depolarization appearing in the blastema of control embryos, and its alteration in embryos which do not regenerate due to inhibition of the three targets.

Example 9

Further Characterization of a Role for Ion Flux During Tail Regeneration

[0454] To further analyze the role for ion flux during tail regeneration, the following experiments are performed. Compounds that selectively inhibited candidate ion transporter proteins are used to evaluate the effect on regeneration, ion flux, pH, and/or membrane potential. Exemplary compounds that can be used include small organic molecules, small inorganic molecules, antisense oligonucleotides (e.g., morpholinos, etc), RNAi constructs, blocking antibodies, and dominant negative constructs (e.g., nucleic acids encoding dominant negative proteins).

[0455] To further analyze the role for ion flux during tail regeneration, expression of known regeneration marker genes will be analyzed. Expression of particular genes uniquely expressed in regenerating tissue (such as pentraxin I, ld1, ld2, hes1, Wnt pathway genes, Notch pathway genes, etc.), as well as markers of undifferentiated cells, proliferating cells, neuronal precursors, and apoptotic cells can be examined. Expression is examined during normal regeneration, as well as in the presence of compounds that specifically perturb (e.g., inhibit or promote) the activity of one or more of

the implicated classes of ion transporter proteins. Additionally, continued examination of ion flux, pH, and membrane potential can be examined.

METHODS: The following methods were used in the experiments performed and discussed in Examples 6-9.

Amputation Protocol:

[0456] Tail amputation technique was performed on wild-type embryos at stage 40, using a single cut with a fresh scalpel blade under a dissecting microscope.

Analysis of Morphology:

[0457] Control and experimental instances of regeneration can be visually evaluate. In some instances, time-lapsed image analysis using a digital camera over several days can be used to provide MPEG movies of normal and altered regeneration. Regardless of whether effects on regeneration are evaluated with the aid of digital photography, statistical analysis of the effects of various treatments on regeneration was and will be performed using two methods. The most basic is a regeneration index: using a constant magnification on a dissecting microscope, embryos undergoing control and experimental regeneration can be photographed, and the analysis tools of the OpenLab software can be used to measure the length of the tails from the end of the hindgut to the caudal tip, along the midline. The degree of regeneration can be quantified as the percent length of the regenerate relative to control. Such indexes can easily be averaged or subjected to ANOVA or other analyses.

Confocal Microscopy of Ion Flux:

[0458] As outlined above, imaging using any of a number of fluorescent dyes including pH- and voltage-sensitive dyes such as SNARF, BCECF, and DiBAC₄(3) can be performed. For specific details regarding use of DiBAC and SERIS, see above methods.

Example 10

Regeneration Studies in Planaria

A Role for Gap Junctional Communication in Anterior Character of Regenerating Planaria

[0459] The restoration of body structures following injury requires an initiation of growth and an imposition of correct morphology upon the regenerating tissue. Understanding this process is crucial both for the basic biology of pattern formation as well as for developing novel biomedical approaches. As used throughout, the term regeneration reflects an appreciation that methods of inducing regeneration (e.g., regeneration of cells or of tissues) require an interplay of proliferative and differentiation events, and may in some tissues also involve dedifferentiation.

[0460] Planaria possess remarkable powers of regeneration. Regeneration is fairly rapid (complete after, 7 days) and is dependent upon a population of stem cells (neoblasts). After bisection across the main body axis, the anterior blastema will regenerate a head while the posterior blastema will regenerate a tail.

[0461] Genes from the family now known as Innexins (formerly called OPUS) comprise a set of important developmental proteins that show no sequence homology to connexins but have the same topology, including four transmembrane domains. The ability of innexins to form functional gap junc-

tion channels has been demonstrated directly for a number of innexins (Landesman et al., 1999, J Cell Sci 112: 2391-2396; Phelan et al., 1998b, Nature 391: 181-184; Stebbings et al., 2000, Mol Biol Cell 11: 2459-2470). Developmental roles of this gene family have been investigated in *Drosophila* and *C. elegans*, where analysis of genetic mutants implicated innexins in the development of muscle and neuronal cell types. However, a role for GJC mediated by innexins during regeneration had not previously been investigated or identified.

[0462] To facilitate analysis of GJC mediated by innexins, we cloned the members of the Innexin family in the planarian, *Dugesia japonica*, and characterized their expression in intact worms and during stages of regeneration. We then performed loss-of-function experiments using compounds that inhibit the expression or activity of multiple innexin family members to test the role of gap junctions in planarian regeneration. The induction of bipolar 2-headed animals following exposure to a GJC blocker that inhibits expression of multiple innexin family members demonstrated a role for GJC during regeneration in planaria.

[0463] Current efforts by the community of planaria researchers have failed to identify connexin genes in the planaria genome. Therefore, we focused our study of a potential role for gap junctional communication during planaria regeneration on another class of gap junctions: the innexins. Innexin genes are known to underlie gap junctional communication in invertebrates many invertebrate species (Dykes et al., 2004, J Neurosci 24: 886-894; Landesman et al., 1999, J Cell Sci 112: 2391-2396; Phelan et al., 1998, Trends in Genetics 14: 348-349; Phelan and Starich, 2001, Bioessays 23: 388-396).

[0464] To isolate planarian innexin genes, we pursued degenerate PCR amplification of innexin gene fragments, using the planarian cDNAs as the templates. We isolated 6 fragments of innexin-like clones, inx1 to 6. We screened a cDNA library to isolate full-length clones. We isolated and sequenced full-length clones for inx1-5. However, inx6 was not present in the cDNA library. We then searched the planarian *D. japonica* EST database and found an additional 7 putative innexins which were present as incomplete fragments. Based upon these, we screened a cDNA library, and isolated full-length cDNA clones, inx7 to inx13.

[0465] All cDNA clones included the initiation codon and the 5' and 3' untranslated sequences. The completed sequences of cDNA clones (inx1-5 and inx7-13) and the sequence of PCR fragment of inx6 have been deposited in the DDBJ/EMBL/GeneBank Library database under accession numbers AB189262, AB189252, AB189253, AB189254, AB189255, AB196957, AB189256, AB189257, AB189258, AB189259, AB189260, AB178521 and AB189261. The foregoing accession numbers and accompanying information is hereby incorporated by reference in their entirety.

[0466] The conserved four transmembrane domains, cysteine residues in the extracellular loops and tetrapeptide sequence (YYQW, located near the end of the first extracellular loop next to the second transmembrane domain), which exist specifically in all innexin sequences reported so far were also present in the planarian innexin sequences, except for *D. japonica* inx1 (because it has a stop codon in the third transmembrane domain). These data indicated that these clones are members of the innexin gene family.

[0467] To gain insight into possible roles of GJC in regeneration, we characterized the expression of innexin genes in the planarian using whole-mount *in situ* hybridization. inx1-

and inx7-positive cells were present throughout the anterior and two posterior branches of the intestine. Expression changed dynamically during regeneration. In head fragments at 2 days after cutting, inx1 and inx7 were expressed in the two small projections corresponding to the early regenerating posterior branches of the intestine. The regenerating branches expressing inx1 and inx7 extended posteriorly and the regenerating pharynx appeared in the anterior region between them at 5 days after cutting. In 1-2 day tail fragments, the intestine branches expressing inx1 and inx7, which had been originally the posterior branches in the intact worms, integrated at an anterior position. inx1 and inx7 were also expressed in one small projection that appeared at the anterior position of the integrated branches, corresponding to the early regenerating anterior branch.

[0468] inx2, inx3, inx4, and inx13 were expressed in the nervous system. Although they were expressed in both the brain and ventral nerve cord (VNC), the distribution of positive cells was different among these genes. inx2 was expressed weakly in the medial region of brain and the medial-distal region of brain branches. inx2 was expressed very weakly in the VNC. inx3 was expressed throughout the brain, and strongly in the medial and lateral regions of the brain and branches. inx3 expression extended to the distal region of the brain branches. inx3 was expressed in the VNC, though the intensity of expression was very low in intact worms, compared to the high expression observed during regeneration.

[0469] inx4 was expressed in the brain branches and the medial and lateral regions of the brain. In contrast to inx2 and inx3, inx4 was expressed in neuron-like cells throughout the peripheral region of the head, where sensory organs are aligned and project to the brain branches. inx4 was expressed in the posterior blastema at 5 days after cutting. inx4 was also expressed in the VNC, and this expression was up-regulated in the anterior region at 5 days after cutting.

[0470] Additionally, inx4 was expressed in a number of cells throughout the body; with especially strong expression detected in the photoreceptor cells that plug the eyecup of the pigment cells. During regeneration, the expression of inx4 in photoreceptor cells begins at 4 days after cutting and prior to the appearance of the maturely pigmented eyecup at 5 days after cutting.

[0471] inx13 was expressed in the lateral and medial region of the brain. In intact worms, the expression in the lateral region was much higher than in the medial region. It was expressed in the brain branches, though the expression was restricted to the stem region. Expression of inx13 in the VNC was very weak, but was up-regulated in the posterior region in the VNC during regeneration.

[0472] In addition to the expression observed in the nervous system, these genes were expressed in the pharynx. inx2, inx3, and inx13 were expressed in the posterior region of the pharynx, while inx4 was expressed in the anterior and posterior regions. Despite expression of these innexin family members in the pharynx, the intestine-type innexins (inx1, inx7) were not expressed in the pharynx.

[0473] The expression of inx2, inx3, inx4 and inx13 changed dynamically during brain regeneration. We categorized dynamic expression into two categories: early (initiating in the regenerating brain within 1 day after cutting) and late (initiating at 2 days after cutting). inx2 and inx4 were late genes. The expression of inx2 was initiated in the medial and lateral region of the regenerating brain at 2 days after cutting.

At 3 days to 4 days after cutting, *inx2* was expressed in the broad region regenerating the brain branches in the anterior blastema. The expression of *inx4* was initiated at the anterior-medial region of the regenerating brain at 2 days after cutting. At 3 to 5 days after cutting, *inx4* was expressed in the medial region of the regenerating brain. The expression of *inx4* was up-regulated transiently in the medial region of the regenerating brain at 4 days after cutting.

[0474] In contrast, *inx3* and *inx13* were early genes. Expression was first detected at 18 hours and 1 day after cutting, respectively. The expression of *inx3* initiated in the early regenerating brain in the anterior region of the blastema within 1 day. The earliest detectable signal was seen at 18 hours after cutting. At 2 days, *inx3* was expressed in the medial and lateral region of the regenerating brain. At 4 to 5 days, the strong expression of *inx3* delineated clearly the structure of brain branches. The expression of *inx13* initiated in the early regenerating brain in the anterior region of the blastema at 1 day after cutting. At 2 days after cutting, *inx13* was expressed in the medial and lateral region of the regenerating brain. At 3 days, *inx3* was expressed in the stems of early regenerating brain branches. At 4 to 5 days after cutting, the expression of *inx13* in the regenerating brain branches grew out peripherally, following the regeneration of the brain branches, but did not extend completely to the tip of the brain branches.

[0475] In intact worms, *inx5* was expressed at the edge of the head where sensory organs are aligning and in the scattered cells distributing throughout the dorsal side of the body. *Inx5* expression exhibiting gradated distribution from the head to tail along the AP axis, and in a number of cells along the VNC, with a dense distribution along the VNC in the head region. During regeneration, *inx5* was expressed in the blastema. At 2 days after cutting, *inx5* was initially expressed at the edge of the anterior blastema and in some scattered blastema cells. Sectioning revealed that *inx5* was expressed at the leading edge of head mesenchyme in the regenerating head. Following brain regeneration, the *inx5*-positive cells appeared at a high density along the VNC in the regenerating head region and in the regenerating tail region.

[0476] *inx12* was expressed very weakly in the head and tail region in intact worms. During regeneration, *inx12* was expressed in both of the anterior and posterior blastema and weakly in the midline in the posterior region of the body. Sectioning revealed that *inx12* was expressed in the mesenchyme anterior to the regenerating intestine in the anterior blastema at 2 days after cutting. At 5 days after cutting, the expression level of *inx12* was reduced in the blastema, and the expression was mostly restricted at the edge of the regenerating head. Following brain regeneration, *inx12* was expressed in cells outlining the VNC in the regenerating head.

[0477] *inx8* and *inx9* were expressed in the mesenchyme throughout most of the body but not in the intestine. *inx8* and *inx9* were expressed in the mesenchyme between the epithelium/muscle, intestine and nervous system, though there were some differences: *inx8* was strongly expressed in some mesenchyme cells around and between the small branches of intestine and between the intestine branch and pharynx, as well as in the pharynx; *inx9* was more ubiquitously expressed in the mesenchyme, but not expressed between the intestine branch and pharynx. Both *inx8* and *inx9* were strongly expressed in the mesenchyme tissue around the pharynx and at the midline in the tail region. Although *inx8* and *inx9* were strongly expressed in the regenerating head and tail at a late

stage of regeneration, *inx9* was highly expressed in the anterior blastema. Sectioning revealed *inx9* in the thin mesenchyme layer outlining the anterior part of the regenerating intestine in the anterior blastema at 2 days after cutting. *inx11* was also expressed in the mesenchyme. Additionally, *inx11* was strongly expressed in the dorsal midline of the body. In contrast to the expression pattern of *inx8* and *inx9*, the expression of *inx11* was restricted to the medial region in the head mesenchyme.

[0478] *inx10* was expressed in a number of small thread-like structures mainly in the lateral-peripheral region in the intact worms. The threadlike structures were sparsely distributed in the mesenchyme tissue underneath the epithelium. This was similar to the known distribution of the protonephridia observed in electron microscopy studies reported previously. During regeneration, the shape of threadlike structures expressing *inx10* changed dynamically in the blastemas. *inx10* was expressed also in the anterior and posterior regions of the pharynx, similarly to the expression as *inx4*.

[0479] To test the hypothesis that gap junctional communication was required for correct patterning during regeneration, we sought an inhibitory compound that would broadly disrupt the function of this class of gap junctions. Currently-popular RNAi approaches are not well-suited for this purpose because they target individual innexin transcripts. Thus, we selected heptanol, a compound that specifically disrupts a broad classes of gap junctions. Heptanol and other long-chain n-alkanols are efficient and rapidly-reversible inhibitors of both electrical and chemical GJC in both connexin and innexin-based gap junctions. In the context of the planarian system, heptanol can be used to disrupt the activity of the broad class of innexin gap junctions.

[0480] Regenerating worms were contacted with the compound at an early stage of regeneration (2 days after cutting). 1-10 μ M heptanol was dissolved in the medium. This concentration was not toxic to the worms, and additionally did not cause observable morphological defects in intact worms. At 7 days post-cutting, we assayed the worms for the morphology of blastemas. Trunk fragments of worms exposed to heptanol exhibited clear anteriorization of both blastemas in 43% of the cases ($n=423$). The range of anteriorized phenotypes included a loss of tail development, ectopic pharynx posterior to the primary pharynx, appearance of an ectopic eye in the posterior blastema, or a complete head at the posterior end (16% for complete bipolar heads); such bipolar anterior (janus) animals were fully viable. Thus, inhibition of this class of ion transporter proteins modulated regeneration in planaria.

[0481] In contrast, all worms regenerating in spring water exhibited normal regeneration ($n=107$). Exposure to hexanol, a reagent similar to heptanol but which is much less effective at blocking GJC, never induced strong anteriorization of the posterior blastema. However, this weaker compound did inhibit tail regeneration (the weakest class of anteriorization). This phenotype is consistent with a dependence of anteriorization upon the degree of GJC inhibition. Importantly, GJC inhibition induced the growth of anterior structures (in many cases, well-formed ectopic heads) and not simply a cessation of regeneration, ruling out toxicity as the mechanism and implicating GJC in events that determine the axial identity of the structure formed during regeneration.

[0482] We next sought to ascertain whether the anteriorizing effect was dependent on the AP (anterior-posterior) level from which the fragment originated. Worms were amputated

at four levels to make five body fragments: head, pre-pharyngeal, trunk (including the pharynx), post-pharyngeal and tail fragments. To enable quantitative analysis of the effect on regeneration, we defined a simple continuous “anteriorizing index” on which each worm was scored as normal or exhibiting weak/strong/complete anteriorization. This allowed a direct comparison of the effects observed in each treated group. To briefly summarize, the strongest anteriorization due to GJC inhibition was observed in the pre-pharyngeal and trunk fragments (anteriorization indexes of 25.8 and 27.6 respectively). The head and post-pharyngeal fragments were less sensitive (anteriorization indexes of 5.6 and 6.2 respectively). No effect was observed on tail fragments. These data are consistent with a role for GJC in mediating the axial patterning along the anterior-posterior axis during regeneration in the planarian.

[0483] To analyze at a molecular level the patterning changes induced in regenerating worms following inhibition of innexin-mediated GJC, we performed whole-mount in situ hybridization analysis of marker genes in bipolar worms. The CNS marker DjPC2 (Agata, 1998, Zoological Science 15: 433-440) was expressed in the brain, VNC and posterior position of the pharynx in the control worms. In perturbed worms exhibiting the bipolar head phenotype, DjPC2 was expressed in the brains (two brains—one brain located at each end) and two pharynxes that lay asymmetrically as mirror images.

[0484] The brain marker DjotxB (Umesono et al., 1999, Dev Genes Evol 209: 31-39) was expressed in the brain and the cells outlining the posterior half of mouth in the control worms. In perturbed worms exhibiting the bipolar head phenotype, DjotxB was expressed in the brains (two brains—one brain located at each end) and in the mirror imaged-mouths.

[0485] The innexin gene, inx7, is a good marker of the intestine. Normally, the intestine has an asymmetric shape along the AP axis: it has one intestine branch anteriorly connected to the pharynx and two intestine branches located posteriorly. In perturbed worms exhibiting the bipolar head phenotype, inx7 expression indicated that the intestine were symmetrically aligned.

[0486] In control, untreated worms, the tail marker DjAbd-Ba (Nogi and Watanabe, 2001, Develop Growth Differentiation 43: 177-184) was expressed strongly in the tail region posteriorly to the pharynx. In perturbed worms exhibiting the bipolar head phenotype, DjAbd-Ba was expressed weakly and broadly in the domain laterally to the pharynxes in the trunk region, and was not expressed in the originally-posterior region in the body.

[0487] These results demonstrated a role for gap junctional communication mediated by the innexin class of ion transporter proteins in regeneration in planaria. Specifically, gap junctional communication modulates regeneration, as well as the anterior character of regenerating fragments.

Worm Husbandry

[0488] The asexual clonal strain GI of the planarian, *Dugesia japonica*, was used in these study. In all experiments, the worms were starved for 1 week before use.

PCR-Based Cloning of the Innexin Genes

[0489] cDNA from regenerating head and tail fragments of planarians (mixed stages at 1-6 days after cutting) were used as templates for PCR to amplify the planarian innexin genes

from a library (5×10^6 independent clones) using the forward primer 5'-CGCGGATCCWSNRRNCARTAYGTNGG-3' and degenerate reverse primer 5'-CGGAATTCGGNAC-CCAYTGRRTARTA-3', corresponding to the highly conserved regions of innexin genes. The amino acid sequences of these highly conserved regions are (S/T)(K/G)QYVG and YYQWVP, respectively. The PCR amplification was carried out with one cycle at 94° C. for 1 min, followed by 40 cycles of 30 sec at 94° C., 30 sec at 45° C. and 30 sec at 72° C., and by a final extension at 72° C. for 5 min. The library was screened by the PCR-based stepwise dilution method (Watanabe et al., 1997, Anal Biochem 252: 213-214).

Whole-Mount In Situ Hybridization

[0490] Whole-mount in situ hybridization was performed as described previously according to standard methods with the following modifications employed for greater sensitivity and lower background. Prior to prehybridization, the samples were incubated twice in 0.1 M triethanolamine, pH 7.6, for 15 min at room temperature, and were acetylated using an acetic anhydride series (0.25% and 0.5%) in 0.1 M triethanolamine, pH 7.6, for 15 min each at room temperature. Hybridization was carried out in hybridization solution (50% formamide, 5×SSC, 100 µg/ml yeast tRNA, 100 µg/ml heparin sodium salt, 0.1% Tween-20, 10 mM DTT, 5% dextran sulfate sodium salt) including about 40 ng/ml digoxigenin (DIG)-labeled antisense riboprobe that had been denatured at 70° C. for 10 mm.

Drug Exposure for GJC Inhibition

[0491] Intact worms 1-1.5 cm long were put into heptanol (or hexanol) solution (0.0045-0.006% vigorously vortexed into spring water) immediately prior to amputation to equilibrate the worms with the drug solution. The worms were amputated at four levels to generate the head, pre-pharyngeal, trunk (or “pharyngeal”), post-pharyngeal, and tail fragments. Worm fragments were incubated at 22° C. for 2 days. The heptanol solution was exchanged for fresh solution every day. The worms were then washed with water twice and incubated in worm water for 14-20 days to monitor the phenotypes.

Scoring System for Anterior-Posterior Phenotype of Regenerates

[0492] We developed a quantitative scheme allowing comparison of degree of anteriorization among groups of worms. Each worm was scored on the following scale by observing the posterior blastema: 0 points—normal (a normal worm with a fully-patterned tail), 1 point—weak anteriorization of posterior blastema (missing tail or bipolar pharynx), 2 points—stronger anteriorization of posterior blastema (incomplete ectopic head with eye structures), or 3 points—complete anteriorization of posterior blastema (bipolar head, where the ectopic head has complete development with 2 normal eyes). For each group of worms, we calculated an average score that is the sum of all scores for the worms divided by the total number of worms. For convenience, the index was scaled from 0 to 100 (final index=average score*100/3). On this scale, a group of worms that were all normal would score 0, while a group of worms all of which were fully double-head would score 100. This scheme was

focused on ascertaining the extent of anteriorization as judged by external morphology.

Example 11

Characterization of the Role of V-ATPase H⁺ Pump in Regeneration

[0493] Further experiments were done in the *Xenopus* system based on the results of the pharmacological screen (see Example 6). In contrast to normal regeneration taking place when larvae are amputated at stage 41 (FIG. 12A,B), exposure to 150 nM concanamycin, a potent and highly specific inhibitor of the V-ATPase H⁺ pump, results in a strong inhibition of regeneration in the absence of general toxicity (N=226, U=11628, Z=11.357; FIG. 12C, Table 3, and see methods). The V-ATPase (Nishi, 2002, Nature Reviews Molecular Cell Biology 3: 94-103) generates strong pH and voltage gradients at the expense of ATP, when expressed in vesicular or cell plasma membranes. Analysis of the localization of activated caspase-3 in control and V-ATPase-inhibited larvae (FIG. 12E,F) revealed that regenerating tails normally possess a small apoptotic cell group, but no significant difference was observed in the degree of apoptosis under V-ATPase inhibition, suggesting that an increase of cell death does not account for this failure to regenerate.

[0494] The ability to target different embryonic regions with early injections of mRNA provided an opportunity to phenocopy the pharmacological phenotype with a molecular loss-of-function construct and also to test the spatial requirements for V-ATPase activity. Using a well-characterized dominant negative V-ATPase E subunit YCHE78 (Lu et al., 2002, Journal of Biological Chemistry 277: 38409-15), we confirmed the same phenotype obtained with concanamycin; YCHE78 misexpression at high levels in the tail (detected by GFP lineage label) prevented regeneration as compared with injected animals not exhibiting YCHE78 expression in the tail (N=66, H=100.232, Q=3.556, p<0.01). These data strongly support the necessity for endogenous V-ATPase function in the tail for regeneration.

[0495] The strong inhibition of regeneration (red bars) by pharmacological V-ATPase blockade was largely prevented (FIG. 12D; white bars=no treatment; green bars=PMA injection only) by misexpression of a concanamycin-insensitive yeast P-type H⁺ pump (PMA1.248, yellow bars) 28 in the tail (301% increase in R1 relative to concanamycin-exposed embryos; N=127, H=81.486, Q=4.672, p<0.01). This rescue experiment demonstrated that it is indeed H⁺ pumping provided by the V-ATPase during normal regeneration that is blocked by concanamycin exposure and normally ensures complete regeneration.

[0496] To test for cell-autonomy of the V-ATPase mechanism in regeneration, we isolated embryos in which the dominant negative V-ATPase construct YCHE78 was localized to somites (confirmed by β-gal lineage label, FIG. 12G). Control regenerates possess a significant muscle component by 96 hours post-fertilization (hpa) as detected by expression of the 12-101 skeletal muscle marker (FIG. 12H). In contrast, regenerating tissue in those few larvae exhibiting a weaker YCHE78 effect, and thus some degree of regeneration, contains little or no detectable muscle marker signal (FIG. 12I). Axon outgrowth patterns in YCHE78 injected animals were, however, completely normal in the partially regenerated tail (FIG. 12J). These data suggest that the V-ATPase activity in

the somites is required for muscle regeneration but not for nerve regeneration, supporting a cell-autonomous role for H⁺ pumping in the somites.

[0497] We next examined the endogenous expression of the V-ATPase in the regeneration bud compared to most targets (FIG. 13A,A'). Expression of the c subunit of the V-ATPase could be detected at the mRNA (FIG. 13B,B') and protein (FIG. 13C,C') levels specifically in the regeneration bud within 6 hours of amputation; other V-ATPase subunits were also up-regulated (data not shown). A low level of background expression elsewhere in the trunk was detected, due to the ubiquitous vacuolar form of the V-ATPase (data not shown). However, the strong plasma membrane expression was observed only in the regeneration bud. Thus, the pump is endogenously expressed in a spatio-temporal pattern consistent with an endogenous role in regeneration.

[0498] We also investigated V-ATPase expression in tails cut during a refractory period, during which *Xenopus* larvae cannot regenerate (Beck et al., 2003, Dev Cell 5: 429-39). Expression was normal in refractory tails (FIG. 13D), suggesting that their inability to regenerate was not due to the failure to turn on V-ATPase expression in the regeneration bud but rather to a post-translational step in V-ATPase function. To determine whether the V-ATPase is normally up-regulated in existing cells or produced by a new cell population generated in response to amputation we irradiated larvae—a procedure known to abolish cell proliferation (Li et al., 2001, Comp Biochem Physiol A Mol Integr Physiol 130: 133-40; Salo and Baguna, 1985, J Embryol Exp Morphol 89: 57-70). Irradiated larvae still up-regulated V-ATPase expression in the wound (FIG. 13E). We confirmed the loss of proliferative cells (FIG. 13F) after X-irradiation, and the resulting failure to regenerate (FIG. 13G). These data suggest that the V-ATPase up-regulation takes place in existing wound cells and does not require the production of a new cell type in the regeneration bud.

[0499] We next directly examined the physiology of the regeneration bud using the voltage reporter dye DiBAC₄(3). Consistent with IHC localization of V-ATPase in the bud (FIG. 13B-C'), we found that in a normal regenerating tail, the bud is depolarized relative to the rest of the tail (FIG. 13H). To determine relative depolarization of uncut versus treated tails, the gain of the photomultiplier tube was reset below that used to generate FIG. 13H. With these adjusted conditions, it is possible to see that control regeneration buds (FIG. 13J) are somewhat depolarized relative to uncut tails (FIG. 13I), while V-ATPase-inhibited (FIG. 13K) are more dramatically depolarized, as evidenced by the higher fluorescence intensity. Refractory tails also exhibited a strong depolarization (FIG. 13L) despite normal expression of the V-ATPase.

[0500] To ask whether the V-ATPase controls regeneration directly via its ion pumping activity, we depolarized tails with a method not relying on V-ATPase: 2 nM palytoxin (Castle and Strichartz, 1988, Toxicon 26: 941-51; Hilgemann, 2003, Proc Natl Acad Sci USA 100: 386-8). This resulted in a 33% reduction of regeneration index (N=81, U=990.5, Z=2.926, p=0.002), suggesting that it is indeed the membrane voltage level (and downstream effects on voltage-sensitive proteins) that is crucial for regeneration. We conclude that consistent with its expression, the ion pumping activity of the V-ATPase is an important determinant of the steady-state membrane polarization level in the regeneration bud cells. Moreover, these data demonstrate that refractory regeneration buds are unable to maintain normal polarization because of a process

downstream of, or distinct from, V-ATPase expression. One possibility is the existence of an as-yet-unidentified depolarizing transporter functioning in refractory tails.

[0501] To determine whether induction of H⁺ flow is a promising strategy for inducing regeneration in a gain-of-function application, we tried to rescue the ability to regenerate during the refractory period by misexpression of the yeast PMA1.2H⁺ pump. Remarkably, expression of PMA1.2 led, in 21/55 refractory tadpoles, to significant regeneration (FIG. 13M,M'); ectopic up-regulation of outgrowth occurring in and at 90° to the neural tube (red arrows); RI increase of 287% over control refractory tails; N=103, U=1638.5, Z=13.005, p<<0.001). This is the first example of the induction of regeneration by molecular expression of an ion transporter and provides a novel entrypoint into this complex process that may be exploited by future clinical augmentation efforts. Consistent also with a general control of growth by H⁺ transport (Cone and Tongier, 1971, Oncology 25: 168-82; Gillies et al., 1992, Cell Physiology and Biochemistry 2: 159-179), misexpression of this ion pump can activate growth even in uncut tissue, and this effect is not restricted to the tail (FIG. 13M").

[0502] To gain insight into the cellular mechanisms by which the V-ATPase participates in regeneration, we characterized the pattern of proliferating cells during regeneration, using an antibody to the phosphorylated Histone 3B—a standard marker of cells in the G₂/M transition of the cell cycle, useful for identifying mitotic cells in regenerating systems including *Xenopus* (Saka and Smith, 2001, Dev Biol 229: 307-18; Sanchez Alvarado, 2003, Current Opinion in Genetics & Development 13: 438-444). At 24 hpa, this subset of proliferating cells is homogeneously distributed throughout the growing tail (FIG. 14A). In contrast, by 48 hpa, these cells are (as expected) highly enriched in the regeneration bud, but, surprisingly, are largely absent from the region of the flank anterior to the amputation (FIG. 14B). Thus, normal regeneration includes two components: an increase in proliferation in the bud, and a >2.3-fold reduction of proliferation in the flank rostral to the amputation plane. The reduction of surrounding proliferating cells does not occur when a puncture wound is made (FIG. 14C; red arrows point to H3P-positive cells [see Methods]; green circle indicates the location of the puncture wound), providing a mechanistic readout of the difference between true regeneration and wound healing. Specific inhibition of the V-ATPase resulted in an approximately 6-fold decrease in the number of proliferating cells in the regeneration bud (FIG. 14D,D',D"), and approximately a 2.5-fold decrease of proliferation in the flank (this reduction did not noticeably impair larval development or behavior). Moreover, V-ATPase-inhibition results in the complete failure of the normal strong reduction of proliferating cells in the flank at 48 hpa (F=11.02, p=0.0002, Table 4). These data suggest that the V-ATPase is required for the up-regulation of proliferation in the growth zone after amputation, as well as for the normal loss of proliferation in the flank. The purpose of this mid-flank reduction is unknown, but it demonstrates that regeneration is not a purely local phenomenon (since mid-flank tissues at a distance of up to 3 mm anterior to the regeneration bud must receive a regeneration-specific signal to stop proliferating) and that such long-range signaling is V-ATPase dependent. Gap junctions are another ion flow control mechanism that has recently been shown to carry long-range information in a complex regenerating system (Nogi and Levin, 2005, Dev Biol 287: 314-35).

[0503] We then sought to functionally link V-ATPase activity to gene expression in the regeneration bud. Existing markers include Notch pathway genes². However, these are only expressed in tissue that does not exist in V-ATPase-inhibited larvae, and thus cannot be examined in V-ATPase loss-of-function regeneration buds. Because of this, and because the V-ATPase is expressed so soon following amputation, we utilized an earlier marker that is normally expressed by 12 hpa: the K⁺ channel KCNK1 (FIG. 14E,E'). In larvae in which V-ATPase activity was abrogated by concanamycin or a function-blocking anti-V-ATPase antibody, KCNK1 expression was absent (N=13, FIG. 14F,F',G,G'). Thus, V-ATPase is upstream of some gene expression in the regeneration bud, including other ion transporters also specifically expressed during early stages of regeneration.

[0504] Finally, to characterize downstream morphogenetic consequences of V-ATPase abrogation, we traced axonal paths during regeneration using immunohistochemistry with an acetylated α-tubulin antibody. In normally regenerating tails, axons appear increased in number (relative to the uncut portion of tail) and they extend into the bud in bundles parallel to the main anterior-posterior axis of the tail (FIG. 14H; green arrows indicate normal axon patterning, while black arrows indicate abnormal axon number and/or location). In contrast, the axons of V-ATPase-inhibited tails increase in density, but axon patterning is abnormal, with axons absent from the middle of the regeneration bud (FIG. 14I) or appearing tangled at the tail tip (FIG. 14I"). These data demonstrate that V-ATPase is required not only for expression of marker genes in the regeneration bud and the increase in proliferation in the growth zone, but also for the patterning of axons in the tail. The known dependence of axon orientation on electrical cues (McCaig et al., 2002, Trends in Neurosciences 25: 354-9) suggests a testable model whereby regeneration currents serve to orient the growth of axons during regeneration of the tail. Consistent with this, our data show that expression of the yeast proton pump, which was able to rescue V-ATPase-inhibited and refractory-inhibited regeneration, also restored normal axon patterning to concanamycin-treated tails (FIG. 14J). Moreover, in normal refractory tails, there is no apparent increase in the number of axons, and those that are present terminate well anterior of the tail tip (FIG. 14K). However, expression of the yeast proton pump PMA in refractory tails induced both the proliferation and axonal patterning normal to regeneration, in tails in which morphological regeneration was and was not induced (FIG. 14L,L'). In those larvae in which normal tail outgrowth was not rescued by PMA, the presence of axons at the very edge of the wound was induced in 25/30 animals (FIG. 14L), demonstrating that the neural patterning and outgrowth are distinct components of the regenerative response, both downstream of H⁺ flux.

[0505] We next sought to probe the relationship between the two major patterning mechanisms observed to depend on the V-ATPase. In order to determine whether the abnormal axonal patterning observed in V-ATPase-inhibited larvae is caused by the inhibition of cell proliferation, we γ-irradiated larvae to abolish proliferation (FIG. 13F). In such animals, despite a lack of regeneration (FIG. 13G), axonal patterning extends all the way to the tip in bundles parallel to the main axis of the bud (FIG. 14M,M'), unlike what we observed due to V-ATPase inhibition (FIG. 14I,I"). Thus, the patterning of axons depends upon V-ATPase activity in a pathway parallel to the induction of cell proliferation, and is not a secondary consequence of mitotic activity.

[0506] Our characterization and loss- and gain-of-function data demonstrate a consistent expression and function of the V-ATPase, and reveal it as a novel biophysical component required for early steps of regeneration of a complex vertebrate appendage. Based on these data, we suggest a model that integrates the known molecular genetic and physiological components (FIG. 15). Amputation triggers a cassette of ion transporter expression in existing cells, with V-ATPase functionally upstream of KCNK1. The activity of the V-ATPase and its downstream transporters results in a specific range of membrane polarization in the regeneration bud cells, leading to an up-regulation of mitosis and axonal outgrowth, ultimately resulting in the regeneration of the tail. Refractory stage larva cannot regenerate due to an extreme depolarization occurring despite normal V-ATPase expression. It is the crucial events downstream of the ion pumping activity of the V-ATPase (occurring as early as 6 hpa) that are disrupted by concanamycin or dominant negative V-ATPase mutant expression. The voltage gradient in the bud depends on the V-ATPase, but the precise physiological state is likely to be a complex function of a module including a number of other ion transporters. Ectopically-induced H⁺ flux can be used to rescue upstream steps and initiate the program of regeneration, representing a tractable “master control” point for therapeutic approaches.

[0507] Comparison of regenerating, refractory, and uncut tails’ voltage maps (FIG. 13H-L) revealed that regenerating buds maintain a moderate level of membrane voltage depolarization, and that deviations from this permissive zone towards either strong depolarization (in refractory, V-ATPase-inhibited) or hyperpolarization (in uncut) is associated with a quiescent, non-growing condition. The permissive physiological state is established by 24 hours, and this level of membrane voltage in regeneration bud cells is not achieved by refractory tails (this physiological property is currently the earliest known mechanistic difference between permissive and refractory stages). Consistent with the known dependence of mitosis on membrane voltage and ion transport activity (Cone, 1974, Annals of the New York Academy of Sciences 238: 420-35; Cone and Cone, 1976, Science 192: 155-8; Olivotto et al., 1996, Bioessays 18: 495-504), this bioelectrical state leads to the required proliferation in the bud and, in parallel, to axonal outgrowth into bud tissues. Galvanotactic guidance is a likely mechanism (Gruler and Nuccitelli, 1991, Cell Motility and the Cytoskeleton 19: 121-133; McCaig et al., 2002, Trends in Neurosciences 25: 354-9) for the effect on neuronal patterning.

METHODS: The following methods were used in the experiments summarized in Example 11.

[0508] Amputation procedure and pharmacological screen *Xenopus laevis* larvae at stage 40-41 (Nieuwkoop and Faber, 1967) had their tails amputated under a dissecting microscope using a scalpel blade at the point where the tail begins to taper. Amputated larvae were cultured in 0.1×MMR/gentamycin drug). Larvae were kept at 22° C. for 7 days and scored for regeneration as below. Drug experiments were carried out at least in duplicate (see Table 2). This screen strategy relies on iterative use of blocker reagents, proceeding from substances of broad targeting to those with high specificity; this results in a binary search that rapidly and inexpensively probes an enormous family tree of all known transporters. Due to the high conservation of ion transporters among phyla, reagents developed in mammalian systems are often useful in invertebrate preparations (Alshuaib and Mathew, 2004, Int J Neuro-

sci 114: 639-50; Carvelli et al., 2004, Proc Natl Acad Sci USA 101: 16046-51; Etter et al., 1999, J Neurochem 72: 318-26; Gasque et al., 2005, J Neurosci 25: 2348-58; Pyza et al., 2004, J Insect Physiol 50: 985-94), and a number of labs have utilized this technique to uncover novel transporters involved in morphogenetic events in both vertebrates and invertebrates (Etter et al., 1999, J Neurochem 72: 318-26; Gasque et al., 2005, J Neurosci 25: 2348-58; Hibino et al., 2006, Development, Genes, and Evolution in press; Pyza et al., 2004, J Insect Physiol 50: 985-94; Shimeld and Levin, 2006, Developmental Dynamics in press). However, because of possible structural divergences among species, the screen results do not conclusively (and we do not claim) that targets not implicated in the screen are not involved. Rather, the screen allowed us to efficiently implicate a small number of specific transporters for molecular validation and characterization.

[0509] Palytoxin (PTX) exposure: palytoxin is a protein from *Palythoa tuberculosa* that converts ubiquitous Na⁺/K⁺ transporters into a non-specific pore leading to rapid depolarization (Castle and Strichartz, 1988, Toxicon 26: 941-51; Hilgemann, 2003, Proc Natl Acad Sci USA 100: 386-8; Tosteson et al., 1997, Ann NY Acad Sci 834: 424-5; Tosteson et al., 2003, J Membr Biol 192: 181-9). We determined that at 2 nM larvae were healthy and behaved normally, despite the inability to regenerate. The penetrance of the regeneration phenotype could be raised by increasing the dose of the PTX, but only at the cost of general toxicity. Thus, in the data described (2 nM), we demonstrate that the regeneration bud is more dependent on membrane voltage level than other cells in the embryo, and we identified this as a dose that could dissociate housekeeping levels from regeneration-specific physiological parameters.

Scoring of Regeneration Efficiency

[0510] To quantify and compare regeneration efficiency of larvae treated with different reagents, we introduced the “Regeneration Index” (RI). Individual larvae within a Petri dish comprising a specific treatment were each scored as follows:

[0511] ++: complete regeneration (regenerated tail, indistinguishable from uncut controls).

[0512] +: robust regeneration in presence of minor defects (missing fin, curved axis).

[0513] +/-: poor regeneration (hypomorphic/defective regenerates).

[0514] -: no regeneration.

[0515] The raw numbers of larvae belonging to each category were calculated; percentages were then multiplied by 3, 2, 1 or 0, for, respectively, ++, +, +/- and -. The RI for that dish, ranges 0-300, with the extreme values corresponding respectively to no regeneration and full regeneration in 100% of the larvae in the sample. The RI evaluates the efficiency of regeneration at the single dish level and allows ready comparison of the effect of treatments to controls.

Statistical Analysis

[0516] To compare among three or more treatments, raw data from the above-described scoring were analyzed using a Kruskal-Wallis test for ordinal data, with H corrected for tied ranks. Post-hoc comparisons were made using Dunn’s Q. To compare between two treatments, raw scoring data were analyzed using a Mann-Whitney U test for ordinal data with tied ranks, and using a normal approximation for large sample sizes. Flank cell data were analyzed using a two-factorial (age and treatment) ANOVA. In all analyses, differences were considered significant if p was less than 0.01.

Proliferating Cell Quantification

[0517] H3P staining marks cells in the G₂/M transition of the cell cycle, and is commonly used for identifying mitotic cells in regenerating systems including *Xenopus* (Saka and Smith, 2001, Dev Biol 229: 307-18; Sanchez Alvarado, 2003, Current Opinion in Genetics & Development 13: 438-444). Fixed specimens at the stages indicated were processed for H3P staining as above using an alkaline-phosphatase secondary. Bleaching of the natural pigments in samples allowed easy counting of H3P-positive cells. The close-up panel to the left illustrates the distinction between purple alkaline-phosphatase signal (H3P-positive nuclei) and brown melanocytes. For quantification, cells were counted manually in the region past the amputation plane. Between 4 and 6 samples were counted for each stage and each condition. Similar numbers were obtained in immunohistochemistry performed on sections as in wholemounts and in fluorescent detection (i.e., reagent penetration and chromogenic staining are not confounding factors in H3P-positive cell detection).

In Situ Hybridization

[0518] Larvae were fixed in MEMFA (Sive et al., 2000) and dehydrated in methanol. In situ hybridization was carried out according to standard protocols (Harland, 1991, *Xenopus laevis*: Practical uses in cell and molecular biology 36: 685-695). The ion transporter constructs used to generate probes for in situ hybridization (ISH) experiments were: Kcnk1 (TWIK-1, BC042262, Open Biosystems), V-ATPase 16 kDa subunit (BE025959, RZPD). Plasmids were used to generate anti-sense riboprobes by in vitro transcription. Experiments included sense probe controls, which exhibited no signal as expected (data not shown). Expression indicated represents consistent consensus patterns obtained from analysis of at least 15 larvae in all experiments.

Immunohistochemistry

[0519] *Xenopus* larvae were fixed overnight in MEMFA, heated for 2 hrs at 65° C. in 50% formamide (to inactivate endogenous alkaline phosphatases; this procedure was not done when using fluorescent secondary antibodies), permeabilized in PBTr+0.1% Triton X100 for 30 min, and processed for immunohistochemistry using alkaline phosphatase secondary antibody (Levin, 2004, Journal of Biochemical and Biophysical Methods 58: 85-96) until signal was optimal and background minimal (usually 12 hrs). Anti-ductin (V-ATPase c' subunit) antibody generated against peptide DAGVRGTAQQPR by Invitrogen reveals 1 single clear band of predicted size on Western blot and was used at 1:500. anti-Caspase-3 (Abcam #AB13847), anti-acetylated α-tubulin (Sigma #T6793), and anti-phospho-H3 (Upstate #05-598) were used at 1:1000. Anti-KCNK1 (a generous gift of Dr. S. A. Goldstein and D. Bockenhauer) was used at 1:500. The 12-101 muscle marker (Gurdon et al., 1985, Cell 41: 913-22; Kintner and Brookes, 1984, Nature 308: 67-9) was used at 1:1 dilution; this monoclonal antibody developed by Jeremy P. Brookes was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biological Sciences, Iowa City, Iowa 52242. Axon and muscle marker detection was performed using fluorescent secondary antibodies. Alexafluor 555-conjugated goat anti-mouse secondary (Invitrogen) was used at 1:500 dilution. Fluorescence images were collected on a Leica TCS

SPZ confocal imaging system; $\lambda_{ex}=543$. Some larvae were embedded in JB4 (Polysciences) and sectioned at 30 μm. Control experiments using no primary antibody and no secondary antibody showed no signal (data not shown). Images in panels H-I of FIG. 12 and in panels H-M' of FIG. 14 were processed using Adobe Photoshop as follows. Background autofluorescence was removed by segmentation of the original fluorescence images. The brightest pixels could readily be selected such that the selection best represented the pattern of axons. This selection was then pasted onto the transmitted light photographs of the same sample. The original un-manipulated images are available on request. Overall image brightness was adjusted for optimal clarity. Images and localization data presented in all figures represent consensus patterns obtained from analysis of at least 15 larvae in all experiments.

Confocal Imaging of Membrane Voltage

[0520] *Xenopus* larvae were soaked in voltage-sensitive dye DiBAC₄(3), (Molecular Probes), at a final concentration of 10 ng/ml in 0.1×MMR in the dark for 30 minutes then imaged with a Leica TCS SP2 Confocal Imaging system, mounted on a Leica upright DM RXE microscope. Because DiBAC₄(3) is anionic, the more depolarized a cell, the greater the accumulation of the permeant dye, and the greater the intensity of intracellular, relative to extracellular, fluorescence. The dye was excited at 488 nm and a 20 nm band of emission wavelengths centered at 515 was collected. If images were to be compared, all images were collected on a single day and photomultiplier gain was kept constant. Fluorescence and transmitted light images were collected. Figures were created using Photoshop™. Transmitted light images may have been manipulated for clarity; fluorescence photos were not manipulated other than by creating overlays, rotating, or cropping.

γ-Irradiation

[0521] Intact larvae were subjected to 10⁴ rads of gamma-irradiation in a Cs¹³⁷ irradiator. The group of larvae was then split in two subgroups, one of which underwent amputation 24 hrs after the irradiation procedure. Larvae were examined periodically and a few of them, for each condition, were fixed at different times after irradiation for immunohistochemistry.

Ion Transporter Misexpression

[0522] Synthetic mRNA was transcribed by the SP6 polymerase from linearized pCS2+plasmids containing the individual cDNAs (YCHE78 and pMA1.2). About 5 ng of each construct mRNA was mixed with 50 ng of RLD and 250 pg of mRNA encoding β-galactosidase, RFP, or GFP (Zernicka-Goetz et al., 1996, Development 122: 3719-24) (as lineage labels) and injected into the 1-cell embryo within 1 hour of fertilization.

[0523] All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

EQUIVALENTS

[0524] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

TABLE 1

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
1-EBIO	K	Ca-activated K channels	Promote
A23187	Ca	Membrane	Inhibit
A23187 (4-Bromo-A23187)	Ca	Membrane	Inhibit
ACA (9-ACA)	Cl	Cl channels	Inhibit
Acetazolamide	H, HCO ₃	Carbonic Anhydrase	Inhibit
Aconitine	Na	Na channel kinetics	Promote
Aflatrem	K	Ca-activated K channels, BK	Inhibit
Agatoxin IIIA (w-Agatoxin IIIA)	Ca	L-, N-, P/Q- and R-type Ca channels	Inhibit
Agatoxin IVA (w-Agatoxin IVA)	Ca	P/Q-type Ca channels	Inhibit
Agitoxin-2	K	Kv channels	Inhibit
Alinidine	K	K-ATP channels	Inhibit
Allethrin	Ca	Cav channels	Inhibit
	Na (to a lesser extent)		
Allethrin	Na	Na channel kinetics	Promote
	Ca (to a lesser extent)		
Allyl isothiocyanate	H, K	H/K-ATPase	Inhibit
AM 92016 HCl	K	K inward rectifier	Inhibit
Amantadine	H	M2 channels	Inhibit
Amiloride	Na, Ca	Na/Ca exchanger	Inhibit
	H (to a lesser extent)		
Amiloride	H, Na	Na/H exchanger (NHE)	Inhibit
	Ca (to a lesser extent)		
Amiloride HCl	Na	Na channels	Inhibit
	Ca and H (to a lesser extent)		
Aminophylline	Cl	CFTR	Promote
Aminopyridine (4-Aminopyridine; 4-AP)	K	Kv channels	Inhibit
Amiodarone	Ca	Cav channels	Promote
Amitriptyline		tight junctions	Inhibit
Amlodipine	Ca	Cav channels	Inhibit
Amlodipine	Ca	N- and P/Q-type Ca channels	Inhibit
Anthranilic acid (AA)	Cl	Cl channels	Inhibit
Apamin	K	Ca-activated K channels, SK	Inhibit
Aprikalim	K	K-ATP channels	Promote
Astemizole	K	KCNH2 (ERG)	Inhibit
Auroverdin B	H	ATP Synthase (F-type H-ATPase)	Inhibit
	K	KCNQ1 (KvLQT1)	Inhibit
Ba ²⁺	K	K channels	Inhibit
	(Na and Ca to a lesser extent)		
Ba ²⁺	Na, Ca	Na/Ca exchanger	Inhibit
	(K to a lesser extent)		
Bafilomycin	H	H-ATPase (V-type H-ATPase)	Inhibit
BAPTA	Ca	Aqueous compartments	Inhibit
Barnidipine	Ca	N- and P/Q-type Ca channels	Inhibit
Batrachotoxin	Na	Na channel kinetics	Promote
Bay K 8644 [(R)-(+)] isomer]	Ca	Cav channels	Inhibit
Bay K 8644 [(S)-(+)] isomer]	Ca	Cav channels	Promote
Benidipine	Ca	Cav channels	Inhibit
Benidipine	Ca	N- and P/Q-type Ca channels	Inhibit

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
Benzamil	Na (H and Ca to a lesser extent)	Na channels	Inhibit
Benzamil	Na, Ca (H to a lesser extent)	Na/Ca exchanger	Inhibit
Benzamil	Na, H (Ca and Na to a lesser extent)	Na/H exchanger (NHE)	Inhibit
Benzamil	Na	Na/HCO ₃ cotransporter	Inhibit
Benzocaine	Na	NaV channels	Inhibit
Benzopyran	K	K-ATP channels	Promote
Benzothiazepines	Ca	Cav channels	Inhibit
BHQ	Ca	SERCA	Inhibit
Bifenthrin	Na	NaV channels	Inhibit
Bimakalim	K	K-ATP channels	Promote
BL-1743	H	M2 channels	Inhibit
BMS-180448	K	K-ATP channels	Promote
Bradykinin	Ca (K to a lesser extent)	Ca stores (internal)	Inhibit
Bradykinin	K (Ca to a lesser extent)	Ca-activated K channels	Promote
Brevetoxin	Na	NaV channels	Inhibit
BRL 55834	K	K-ATP channels	Promote
Burnetanide	Na, K, Cl	Na/K/Cl cotransporter	Inhibit
Bupivacaine	K	K-ATP channels	Inhibit
Butanedione monoxime (2,3-	Ca	SERCA	Promote
Butanedione monoxime)			
Caffeine	Ca	SERCA	Inhibit
Calciseptine	Ca	L-type Ca channels	Inhibit
cAMP	Ca	SERCA	Promote
Carbenoxolone		Gap Junctions	Inhibit
Carbocyanine	K	Ca-activated K channels, IK	Inhibit
Cariporide (HOE642)	Na, H	Na/H antiporter, (NHE)	Inhibit
Cassigarol A	H, K	H/K-ATPase	Inhibit
Catechins	H, K	H/K-ATPase	Inhibit
CCCP	H	Membrane	Inhibit
Cd ²⁺	Cl (C, K, Na to a lesser extent)	Ca-activated Cl channels	Inhibit
Cd ²⁺	K (Ca, Cl, Na to a lesser extent)	KCNQ1 (KvLQT1)	Inhibit
Cd ²⁺	Ca (Cl, K, Na to a lesser extent)	L- and N-Type Ca channels	Inhibit
Cd ²⁺	Na, Ca (Cl, K to a lesser extent)	Na/Ca exchanger	Both - depending on the circumstances
Cefadroxil	H	H/peptide transport	Inhibit
CGP 37157	Na, Ca	Na/Ca exchanger	Inhibit
Chalcone	H, K	H/K-ATPase	Inhibit
Charybdotoxin	K	Ca-activated K channels, BK	Inhibit
Charybdotoxin	K	Ca-activated K channels, IK	Inhibit
Charybdotoxin	K	Kv channels	Inhibit
Chloroquine	K (Tight junctions to a lesser extent)	K-ATP channels	Inhibit
Chlorpropamide	K	K-ATP channels	Inhibit
Chromanol 293B	K (Cl to a lesser extent)	KCNQ1 (KvLQT1)	Inhibit
Ciclazindol	K	K-ATP channels	Inhibit

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
Ciguatoxin	Na	Na channel kinetics	Promote
Cilnidipine	Ca	Cav channels	Inhibit
Cilnidipine	Ca	N- and P/Q-type Ca channels	Inhibit
Cismethrin	Na	NaV channels	Inhibit
Clofilium	K	KCNH2 (ERG)	Inhibit
Clofilium	K	KCNQ1 (KvLQT1)	Inhibit
Clotrimazole	K	Ca-activated K channels, IK	Inhibit
Clotrimazole	K	KCNQ1 (KvLQT1)	Inhibit
Co ²⁺	Cl	Ca-activated Cl channels	Inhibit
Co ²⁺	Ca	L- and N-Type Ca channels	Inhibit
Co ²⁺	Na/Ca	Na/Ca exchanger	Inhibit
Concanamycin	H	H-ATPase (V-type H-ATPase)	Inhibit
Conotoxin GVIA (<i>w</i> -Conotoxin GVIA)	Ca	N-type Ca channels	Inhibit
Conotoxin MVIIIC (<i>w</i> -Conotoxin MVIIIC)	Ca	N- and P/Q-type Ca channels	Inhibit
Conotoxin PVIIA (<i>k</i> -Conotoxin PVIIA)	K	Kv channels	Inhibit
Conotoxin SMIIIA (<i>m</i> -Conotoxin SMIIIA)	Na	Na channels	Inhibit
Conotoxin ViTx	K	K channels	Inhibit
Correolide	K	Kv channels	Inhibit
CP 339818 HCl	K	Kv channels	Inhibit
Cromakalim	K	K-ATP channels	Promote
Cs ⁺	K	K channels	Inhibit
Cs ⁺	K	Kv channels	Inhibit
Cu ²⁺		Aquaporin-3	Inhibit
Cu ²⁺	Na, Ca	Na/Ca exchanger	Inhibit
Cu ²⁺	Na, K	Na/K-ATPase	Inhibit
Cu ²⁺		tight junctions	Promote
CyanoGuanidine	K	K-ATP channels	Promote
Cyclopiazonic Acid	Ca	SERCA	Inhibit
Cyfluthrin	Na	NaV channels	Inhibit
Cyhalothrin	Na	NaV channels	Inhibit
Cypermethrin	Na	NaV channels	Inhibit
Cyphenothrin	Na	Na channel kinetics	Inhibit
Darodipine	Ca	Cav channels	Inhibit
DCCD	H	H-ATPase (V-type H-ATPase)	Inhibit
Decanoic acid	K	K-ATP channels	Inhibit
Decarbamylsaxitoxin	Na	Na channels	Inhibit
Deltamethrin	Na	Na channel kinetics	Inhibit
Dendrotoxin (a Dendrotoxin)	K	Kv channels	Inhibit
Dendrotoxin (b Dendrotoxin)	K	Kv channels	Inhibit
Dendrotoxin (d Dendrotoxin)	K	K inward rectifier	Inhibit
Dequalinium dichloride	K	Ca-activated K channels, SK	Inhibit
DHS-1 (Dehydrosoyasaponin-1)	K	Ca-activated K channels, BK	Promote
Diazepam	Cl	Cl channels	Promote
Diazoxide	K	K-ATP channels	Promote
Dibucaine	K	K-ATP channels	Inhibit
Dichlorphenamide	H, HCO ₃	Carbonic Anhydrase	Inhibit
Dideoxyforskolin	Cl	Cl channels	Inhibit
DIDS	Cl	Cl channels	Inhibit
	(H, K to a lesser extent)		

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
DIDS	K (Cl, H to a lesser extent)	KCNQ1 (KvLQT1)	Promote
Diethylstilbestrol (DES)	H	ATP Synthase (F-type H-ATPase)	Inhibit
Diethylstilbestrol (DES)	H	H-ATPase (V-type H-ATPase)	Inhibit
Dihydropyridines	Ca	Cav channels	Inhibit
Diltiazem (D-cis Diltiazem-HCl)	Ca	Cav channels	Inhibit
Diltiazem (L-cis-diltiazem)	Ca	CNG channels	
Diltiazem (L-cis-diltiazem)	K	Cav channels	Inhibit
Dimethadione	K	K channels	Inhibit
Dimethylamiloride	Na, Ca (H to a lesser extent)	Na/Ca exchanger	Inhibit
Dimethylamiloride	Na, H (Ca to a lesser extent)	Na/H antiporter, (NHE)	Inhibit
Diphenylbutylpiperidine (DPBP)	Ca	L- and T-type Ca Channels	Inhibit
Dofetilide	K	KCNH2 (ERG)	Inhibit
DPMX	Cl	CFTR	Promote
E-4031	K	KCNH2 (ERG)	Inhibit
EDTA	Ca	Aqueous compartments	Inhibit
Efaroxan	K	K-ATP channels	Inhibit
EGTA	Ca	Aqueous compartments	Inhibit
Eicosapentanoic acid (EPA)		tight junctions	Inhibit
EIPA (Ethylisopropyl amiloride)	Na, H	Na/H antiporter, (NHE)	Inhibit
Ellagic acid	H, K	H/K-ATPase	Inhibit
Englitazone	K	K-ATP channels	Inhibit
Enprofylline	Cl	CFTR	Promote
Ergotoxin	K	KCNH2 (ERG)	Inhibit
Erythrosin 5'-isothiocyanate	Ca	SERCA	Inhibit
Erythrosine B	H	H-ATPase (V-type H-ATPase)	Inhibit
Esomeprazole	H, K	H/K-ATPase	Inhibit
Estradiol (11-b-Estradiol)	K	KCNQ1 (KvLQT1)	Inhibit
Estradiol (17-b-Estradiol)	Na, H	NHE	Inhibit
Ethacrynic acid	Cl	Cl channels	Inhibit
Ethacrynic acid	Cl	Cl-ATPase	Inhibit
Ethoxyzolamide	H	Carbonic Anhydrase	Inhibit
FCCP	H	Membrane	Inhibit
Fe ²⁺	Na, Ca (K to a lesser extent)	Na/Ca exchanger	Inhibit
Fe ²⁺	Na, K (Ca to a lesser extent)	Na/K-ATPase	Inhibit
Felodipine	Ca	Cav channels	Inhibit
Fenofibrate	K	K-ATP channels	Inhibit
Fenpropathrin	NA	NaV channels	Inhibit
Fenvalerate	Na	NaV channels	Inhibit
Flecainide acetate	Na	Na channels	Inhibit
Flunarizine dihydrochloride	Ca (Na to a lesser extent)	Ca channels	Inhibit
Flunarizine dihydrochloride	Na (Ca to a lesser extent)	Na channels	Inhibit
FPL 64176	Ca	Cav channels	Promote

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
FR177995	H	H-ATPase (V-type H-ATPase)	Inhibit
Furosemide	Na, K, Cl	Na/K/Cl cotransporter	Inhibit
Fusicoccin	H, K	H-ATPase (V-type H-ATPase)	Promote
Gabapentin HCl	Ca	Cav channels	Inhibit
Gaboon viper venom	K	K inward rectifier	Inhibit
Gadolinium	Ca	Ca-activated Cl channels	Inhibit
Gadolinium		Mechano-sensitive channels	Inhibit
Gallopamil	Ca	Ca-ATPase	Inhibit
Glibenclamide	Ca (K to a lesser extent)	CFTR	Inhibit
Glibenclamide (Glyburide)	K (Ca to a lesser extent)	K-ATP channels	Inhibit
Glipizide	K	K-ATP channels	Inhibit
Glycyrrhetic acid (18-b-Glycyrrhetic acid)	GJC	Gap Junctions	Inhibit
Gonyautoxin II	Na	Na channels	Inhibit
Gonyautoxin III	Na	Na channels	Inhibit
Gramicidin A	K	Membrane	
Grayanotoxin	Na	Na channel kinetics	Promote
Guanethidine	K	K-ATP channels	Inhibit
Guanidine methanesulfonate	Na, H	NHE	Inhibit
Haloperidol	K	KCNH2 (ERG)	Inhibit
Halothane	K	Ca-activated K channels	Inhibit
Halothane	K	K inward rectifier	Inhibit
Halothane	K	K Tandem Pore channels	Promote
Hanatoxin-1	K	Kv channels	Inhibit
Hanatoxin-2	K	Kv channels	Inhibit
Heptanol	GJC	Gap Junctions	Inhibit
HMR-1098	K	K-ATP channels	Inhibit
HMR-1556	K	KCNQ1 (KvLQT1)	Inhibit
HMR-1556	Ca	Cav channels	Inhibit
HOCl		tight junctions	Inhibit
HOE642 (cariporide)	Na, H	NHE	Inhibit
HOE694	Na, H	NHE	Inhibit
Hongotoxin	K	Kv channels	Inhibit
Hydroxydecanoate (5-Hydroxydecanoate)	K	K-ATP channels (mitochondrial)	Inhibit
Hydroxyzine	K	KCNH2 (ERG)	Inhibit
Iberiotoxin	K	Ca-activated K channels, BK	Inhibit
IBMX	Cl	CFTR	Promote
Ibutilide	K	KCNH2 (ERG)	Inhibit
indanyloxyacetic acid (IAA)	Cl	Cl channels	Inhibit
Iodo-resiniferatoxin	Ca	TRPV1	Inhibit
Ionomycin	Ca	Membrane	
Isoflurane	K	K Tandem Pore channels	Promote
Isradipine	Ca	Cav channels	Inhibit
Kaliotoxin	K	Ca-activated K channels, BK	Inhibit
Kaliotoxin	K	Kv channels	Inhibit
KB-R7943 mesylate	Na, Ca	Na/Ca exchanger	Inhibit
Kurtoxin	Ca	T-type Ca channels	Inhibit
Lamotrigine	Ca (Na, K to a lesser extent)	Ca channels	

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
Lamotrigine	K (Na, Ca to a lesser extent)	KCNH2 (ERG)	Inhibit
Lamotrigine	K (Na, Ca to a lesser extent)	Kv channels	Promote
Lamotrigine	Na (Ca, K to a lesser extent)	NaV channels	Inhibit
Lansoprazole	H, K	H/K-ATPase	Inhibit
Lanthanum	Ca	Ca channels	Inhibit
Lanthanum	Cl	Ca-activated Cl channels	Inhibit
Levocromakalim	K	K-ATP channels	Promote
Lidocaine	K	K-ATP channels	Inhibit
Lidocaine n-ethyl bromide	Na	Na channels	Inhibit
Lindane		Gap Junctions	Inhibit
Linoleoylamide	K	KCNH2 (ERG)	Inhibit
Linopirdine	K	Ca-activated K channels	Inhibit
Linopirdine (specific to 2/3)	K	KCNQ2/3	Inhibit
Linopirdine dihydrochloride	K	Kv channels	Inhibit
Lobatamide C	H	H-ATPase (V-type H-ATPase)	Inhibit
Lobatamide C A15	H	H-ATPase (V-type H-ATPase)	Inhibit
Lobatamide C A16	H	H-ATPase (V-type H-ATPase)	Inhibit
Lobatamide C A6	H	H-ATPase (V-type H-ATPase)	Inhibit
Lonidamine	Cl	CFTR	Inhibit
Loperamide	Ca	Cav channels	Inhibit
LY83583	Ca	CNG channels	Inhibit
LY97241	K	KCNH2 (ERG)	Inhibit
Margatoxin	K	Kv channels	Inhibit
MaxiPost	K	Ca-activated K channels, BK	Promote
Mefenamic acid	K	KCNQ1 (KvLQT1)	Promote
Mefloquine	K	K-ATP channels	Inhibit
Meglitinide	K	K-ATP channels	Inhibit
Mepivacaine	K	K-ATP channels	Inhibit
Methazolamide	H, HCO ₃	Carbonic Anhydrase	Inhibit
Methoxy-verapamil	Ca	Cav channels	Inhibit
Mexitetine	K	K-ATP channels	Inhibit
Mg ²⁺	Na, Ca	Na/Ca exchanger	Inhibit
Mibepridil	Ca	L- and T-type Ca channels	Inhibit
Minoxidil	K	K-ATP channels	Promote
MK-499	K	KCNH2 (ERG)	Inhibit
Mn ²⁺	Cl (Na, Ca to a lesser extent)	Ca-activated Cl channels	Inhibit
Mn ²⁺	Ca (Na, Cl to a lesser extent)	Na/Ca exchanger	Inhibit
Mn ²⁺	Na (Ca, Cl to a lesser extent)	Na/Ca exchanger	Inhibit
Moclobemide	H	Aqueous compartments	Promote
Monensin	Na	Membrane	Inhibit
Nateglinide	K	K-ATP channels	Inhibit
NBD (chloronitrobenzoxadiazole)	H	ATP Synthase (F-type H-ATPase)	Inhibit
NBD (chloronitrobenzoxadiazole)	H	H-ATPase (V-type H-ATPase)	Inhibit
Neosaxitoxin	Na	Na channels	Inhibit
N-Ethylmaleimide (NEM)	H	H-ATPase (V-type H-ATPase)	Inhibit

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
Nicardipine	Ca	L-type Ca channels	Inhibit
Nicardipine	Ca	N- and P/Q-type Ca channels	Inhibit
NiCl ₂	Na, Ca	Na/Ca exchanger	Inhibit
NiCl ₂	Ca	T-type Ca channels	Inhibit
Nicorandil	K	K-ATP channels (mitochondrial)	Promote
Nifedipine		KCNG channels	Inhibit
Nifedipine	Ca	Cav channels	Inhibit
Niflumic acid	Cl	Ca-activated Cl channels	Inhibit
Nigericin	K	Membrane	
Niguldipine HCl [(R)(-) isomer]	Ca	Cav channels	Inhibit
Niguldipine HCl [(S)(+) isomer]	Ca	Cav channels	Inhibit
Nilvadipine	Ca	Cav channels	Inhibit
Nimodipine	Ca	Cav channels	Inhibit
Nisoldipine	Ca	Cav channels	Inhibit
Nitrendipine	K (Ca to a lesser extent)	Ca-activated K channels, IK	Inhibit
Nitrendipine	Ca (K to a lesser extent)	Cav channels	Inhibit
Norepinephrine	Ca	Ca-ATPase	Promote
Noxiustoxin	K	Kv channels	Inhibit
NPPB (PAA)	Cl	Cl channels	Inhibit
NS004	K	Ca-activated K channels, BK	Promote
NS1608	K	Ca-activated K channels, BK	Promote
NS1619	K	Ca-activated K channels, BK	Promote
Ochratoxin A	Ca	Ca-ATPase	Inhibit
Oleamide	K	KCNH2 (ERG)	Inhibit
Oligomycin	H (Na, K to a lesser extent)	ATP Synthase (F-type H-ATPase)	Inhibit
Oligomycin	H (Na, K to a lesser extent)	H-ATPase (V-type H-ATPase)	Inhibit
Oligomycin	Na, K (H to a lesser extent)	Na/K-ATPase	Inhibit
Omeprazole	H, K	H/K-ATPase	Inhibit
Orthovanadate	Cl (Na, K, H to a lesser extent)	Cl transport	Inhibit
Ouabain	Na, K	Na/K-ATPase	Inhibit
P1075	K	K-ATP channels (sarcolemmal)	Promote
PAA (NPPB)	Cl	Cl channels	Inhibit
Palytoxin (PTX)	Na, K	Na/K-ATPase	Inhibit
Pantoprazole	H	H/K-ATPase	Inhibit
Paspalicine	K	Ca-activated K channels, BK	Inhibit
Paspalinine	K	Ca-activated K channels, BK	Inhibit
Paspalitrem A	K	Ca-activated K channels, BK	Inhibit
Paspalitrem C	K	Ca-activated K channels, BK	Inhibit
Paxilline	K	Ca-activated K channels, BK	Inhibit
Penfluridol	Ca	T-type Ca channels	
Penitrem A	K	Ca-activated K channels, BK	Inhibit
Pentoxifylline	Cl	CFTR	Promote
pentylenetetrazol	K	K-ATP channels	Promote

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
pentylenetetrazol	K	Kv channels	Both - depending on the circumstances
Permethrin	Na	NaV channels	Inhibit
Phenothrin	Na	Na channel kinetics	Inhibit
Phentolamine	K	K-ATP channels	Inhibit
Phenylalkylamines	Ca	Cav channels	Inhibit
PHM (phentolamine mesylate)	K	K-ATP channels	Inhibit
PI(4,5)P2	K	K inward rectifier	Promote
PI(4,5)P2	K	K-ATP channels	Inhibit
PI(4,5)P2	K	KCNH2 (ERG)	Promote
Pinacidil	K	K-ATP channels	Promote
PKF 217-744	K	K-ATP channels	Promote
Pompilidotoxin (b- Pompilidotoxin; wasp toxin)	Na	Na channels	Promote
Prenylamine	Ca	Ca-ATPase	Inhibit
Procaine	K (Na to a lesser extent)	K-ATP channels	Inhibit
Procaine	Na (K to a lesser extent)	NaV channels	Inhibit
Prodigiosin	H	H/K-ATPase	Inhibit
Prodigiosin	H	H-ATPase (V-type H-ATPase)	Inhibit
Propranolol	K	K-ATP channels	Inhibit
Prostaglandin J2 (15-deoxy- Delta12,14- Prostaglandin J2)	K	K-ATP channels	Inhibit
Pumaprazole	H	H/K-ATPase	Inhibit
Pyrethroids	Na	Na channel kinetics	Inhibit
Quercetin	H, K	H/K-ATPase	Inhibit
Quinidine	K	KCNH2 (ERG)	Inhibit
Quinine	K	K-ATP channels	Inhibit
QX 222	Na	Na channels	Inhibit
QX 314	Na	NaV channels	Inhibit
Rabeprazole	H, K	H/K-ATPase	Inhibit
repaglinide	K	K-ATP channels	Inhibit
Retigabine	K	KCNQ	Promote
Rilmakalim	K	K-ATP channels	Promote
Riluzole	Ca (Na to a lesser extent)	Ca channels	Inhibit
Riluzole HCl	Na (Ca to a lesser extent)	Na channels	Inhibit
Riodipine (Ryosidine)	Ca	Cav channels	Inhibit
RP 66471	K	K-ATP channels	Promote
rTarnapin	K	KCNN2	Inhibit
Ruthenium red	Ca	Ca translocators (many)	Inhibit
Saxitoxin (IIa (—OH) STX)	Na	Na channels	Inhibit
Saxitoxin (PSP)	Na	Na channels	Inhibit
Saxitoxinol (a- saxitoxinol)	Na	Na channels	Inhibit
Saxitoxinol (b- saxitoxinol)	Na	Na channels	Inhibit
SB-242784	H	H-ATPase (V-type H-ATPase)	Inhibit
SCA40	K	Ca-activated K channels, BK	Promote
SCH28080	H, K	H/K-ATPase	Inhibit
Scorpion toxins (American b)	Na	Na channel kinetics	Promote
Scorpion toxins (North African a)	Na	Na channel kinetics	Promote

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
Syllatoxin	K	Ca-activated K channels, SK	Inhibit
Sea Anemone toxins	Na	Na channel kinetics	Promote
Sevoflurane	K (Ca to a lesser extent)	Ca-activated K channels, BK	Promote
Sevoflurane	K (Ca to a lesser extent)	K Tandem Pore channels	Promote
Sevoflurane	Ca	Cav channels	Inhibit
SG 209	K	K channels	Promote
SNX-111	Ca	N-type Ca channels	Inhibit
SNX-482	Ca	R-type Ca channels	Inhibit
sodium caprate		tight junctions	Inhibit
Sophalcone	H, K	H/K-ATPase	Inhibit
Sophoradin	H, K	H/K-ATPase	Inhibit
Sotalol	K	KCNH2 (ERG)	Inhibit
SR 33805 oxalate	Ca	Cav channels	Inhibit
Sr ²⁺	Na, Ca	Na/Ca exchanger	Inhibit
Streptolysin O		Membrane	Inhibit
Strophantidin	Na, K	Na/K-ATPase	Inhibit
TAC-101		Gap Junctions	Promote
Tamoxifen	Cl	Cl channels	Inhibit
TEA	K	Ca-activated K channels, BK	Inhibit
TEA	K	KCNQ	Inhibit
TEA (Tetraethylammonium)	K	Kv channels	Inhibit
Tedisamil	K	K channels	Inhibit
Tefluthrin	Na	NaV channels	Neg
Tenidap	K	K inward rectifier	Promote
Tenidap	Na	Na/HCO ₃ co-transporter	Inhibit
Terfenadine	K	KCNH2 (ERG)	Inhibit
Terikalant	K	K inward rectifier	Inhibit
Teripapin-Q (bee venom toxin)	K	K inward rectifier	Inhibit
Tetracaine		CNG channels	Inhibit
Tetracaine	K	K-ATP channels	Inhibit
Tetrahydroberberine (THB)	K	Kv channels	Inhibit
Tetramethrin	Na	Na channel kinetics	
Tetrandrine	K	Ca-activated K channels, BK	Inhibit
Tetrodotoxin (TTX)	Na	Na channels	Inhibit
Thapsigargin	Ca	SERCA	Inhibit
Theophylline	Cl	CFTR	Promote
Tiludronate (skelid)	H	H-ATPase (V-type H-ATPase)	Inhibit
Tityustoxin-Ka	K	Kv channels	Inhibit
TMB-8	K	K-ATP channels	Inhibit
Tolazamide	K	K-ATP channels	Inhibit
Tolbutamide	K	K-ATP channels	Inhibit
TPEN	Zn, Cu, Fe	Other	Inhibit
TRAM-34	K	Ca-activated K channels, IK	Inhibit
Tributyl Tin (TBT)	Cl, OH	Membrane	
Tricaine	Na	NaV channels	Inhibit
Trifluoperazine	Ca	Calmodulin	Inhibit
Trimethadione	K	K inward rectifier	Inhibit
Troglitazone	K	K-ATP channels	Inhibit
U-37883	K	K-ATP channels	Inhibit
U-37883A	K	K-ATP channels	Inhibit
UCL 1684	K	Ca-activated K channels, SK	Inhibit
Valinomycin	H, K	Membrane	
Vanadate	Na, K (Cl and H to a lesser extent)	Na/K-ATPase	Inhibit
Verapamil HCl	K	KCNH2 (ERG)	Inhibit

TABLE 1-continued

COMPOUND	Ion Species	Target ion transporter or class of transporters	Inhibit/promote activity of the target
Verapamil HCl	Ca (K to a lesser extent)	Cav channels	Inhibit
Veratridine	Na	Na channel kinetics	Promote
Verruculogen	K	Ca-activated K channels, BK	Inhibit
Vinpocetine	Na	NaV channels	Inhibit
Way-123398	K	KCNH2 (ERG)	Inhibit
WIN 17317-3	K	Kv channels	Inhibit
Xanthoangelol	H, K	H/K-ATPase	Inhibit
XE 991 dihydrochloride	K	KCNQ1 (KvLQT1), KCNQ1 (KvLQT1) + minK, KCNQ2 + 3	Inhibit
Y-26762	K	K-ATP channels	Promote
YS-035	K	K channels	Inhibit
Zatebradine	K	Kv channels	Inhibit
ZD6169	K	K-ATP channels	Promote
ZD7288	K (Na to a lesser extent)	HCN channels	Inhibit
ZD7288	Na (K to a lesser extent)	HCN channels	Inhibit
ZD7288	Ca	T-type Ca channels	Inhibit
Zetekitoxin AB	Na	NaV channels	Inhibit
ZM 181,037	K	K-ATP channels	Inhibit
ZM 226600 (Anilide tertiary carbinols)	K	K-ATP channels	Promote
Zn	H, Na	NHE	Inhibit
ZnCl ₂	Na, Ca	Na/Ca exchanger	Inhibit

TABLE 2

Compound	Concentration	Target and function	Reference
Amiloride	2.5 mM	Inhibitor of Na ⁺ /H ⁺ antipporter	(Harris and Fliegel, 1999, International Journal of Molecular Medicine 3: 315-21)
4-Aminopiridine	2.6 mM	Blocker of Kv channels	(Abraham et al., 2003, Acta Biol Hung 54: 63-78)
9-anthracene-carboxylic acid	2.5 μM	Inhibitor of Cl ⁻ channels	(Yarar et al., 2001, J Soc Gynecol Investig 8: 206-9)
Benzamil	10 μM	Inhibitor of epithelial sodium channels	(Taguchi et al., 2005, Biochem Biophys Res Commun 327: 915-9)
Diazoxide	10 μM	Opener of K ⁺ _{ATP} channels	(D'Hahan et al., 1999, Proceedings of the National Academy of Sciences of the United States of America 96: 12162-7)
EIPA	50 μM	Inhibitor of Na ⁺ /H ⁺ antipporter	(Pizzonia et al., 1996, Journal of Neuroscience Research 44: 191-8)
EM12	0.7 mM	Inducer of gap-junctional communication	(Onat et al., 2001, Biochem Pharmacol 62: 1081-6)

TABLE 2-continued

Compound	Concentration	Target and function	Reference
Gadolinium chloride	10 µM	Blocker of mechano-sensitive channels	(Krasznai et al., 2003, <i>Cell Motil Cytoskeleton</i> 55: 232-43)
Glibenclamide	1.44 mM	Inhibitor of K ⁺ _{ATP} channels	(Quayle et al., 1997, <i>Physiol Rev</i> 77: 1165-232)
18-β-Glycyrrhetic acid	26.5 µM	Blocker of gap-junctional communication	(Davidson and Baumgarten, 1988, <i>Journal of Pharmacology & Experimental Therapeutics</i> 246: 1104-7)
Heptanol	Dilution of 1 × 10 ⁻⁵	Blocker of gap-junctional communication	(Deleze and Herve, 1983, <i>J Membr Biol</i> 74: 203-15; Takens-Kwak et al., 1992, <i>American Journal of Physiology</i> 262: C1531-8)
Lanthanum chloride	10 µM	Blocker of Ca ⁺⁺ channels	(Nathan et al., 1988, <i>Journal of General Physiology</i> 91: 549-72)
Loperamide	0.2 mM	Ca ⁺⁺ channel blocker	(Harper et al., 1997, <i>Proc Natl Acad Sci USA</i> 94: 14912-7)
Ouabain	16 µM	Inhibitor of Na ⁺ /K ⁺ -ATPase	(Liu, 2005, <i>Front Biosci</i> 10: 2056-63)
PPADS	50 µM	PX27 channels blocker	(Ziganshin et al., 1993, <i>Br J Pharmacol</i> 110: 1491-5)
Quinidine	50 µM	Blocker of slow delayed K ⁺ rectifier	(Yao et al., 1996, <i>J Pharmacol Exp Ther</i> 279: 856-64)
SCH28080	0.12 mM	Inhibitor of H ⁺ /K ⁺ -ATPase	(Vagin et al., 2002, <i>Biochemistry</i> 41: 12755-62)
Suramin	100 µM	Blocker of Cl ⁻ , Ca ⁺⁺ channels	(Bachmann et al., 1999, <i>Naunyn Schmiedebergs Arch Pharmacol</i> 360: 473-6; Emmick et al., 1994, <i>J Pharmacol Exp Ther</i> 269: 717-24)
Tetraethylammonium	1 mM	Blocker of Ca ⁺⁺ -activated K ⁺ channels	(Shen et al., 1994, <i>Pflugers Arch</i> 426: 440-5)
THB	7.3 mM	Inhibitor of voltage-gated K ⁺ channels	(Wu and Jin, 1996, <i>Neurosci Lett</i> 207: 155-8)
Concanamycin	150 nM	Inhibitor of V-ATPase	(Drose et al., 1993, <i>Biochemistry</i> 32: 3902-6; Woo et al., 1996, <i>Biological & Pharmaceutical Bulletin</i> 19: 297-9)

Description of Table 2: Table of Reagents.

[0525] This table lists compounds that were tested for their ability to specifically inhibit regeneration while permitting normal primary tail development, wound healing, and general embryogenesis. The particular concentration of compound used in this study are provided. However, concentration used may vary depending on the biological process and organisms

being evaluated. The major targets are listed for each drug, but many of these compounds also interact with other transporters. Any target of a reagent is thus ruled out of further consideration when a given compound does not affect regeneration, and in these cases broader specificity is a benefit because it allows a greater number of candidates to be filtered out. This screen implicated the V-ATPase transporter that we subsequently validated molecularly and characterized.

TABLE 3

Experiment	N	Regeneration Index		Treated/ Statistical					
		Control	Treated	Control	Test	Statistic	Probability		
Regeneration in Concanamycin PMA-expressing	226	216	49	23%	MWU	U = 11.628	Z = 11.357	p << 0.001	
Regeneration in Concanamycin	127	75	226	301%	KW + DQ	H = 81.486	Q = 4.672	p < 0.01	
Regeneration in Palytoxin	81	187	144	77%	MWU	U = 990.5	Z = 2.926	p = 0.002	
YC7E78-expressing	66	265	194	73%	KW + DQ	H = 100.232	Q = 3.556	p < 0.01	
Regeneration PMA-expressing	103	15	42	280%	MWU	U = 1638.5	Z = 13.005	p << 0.001	
Regeneration of Refractory Tail									

Description of Table 3: Primary Data and Statistical Analysis for Regeneration Assays

[0526] Effect of the inhibitor of V-ATPase on tail regeneration is expressed as a "Regeneration Index" (RI) computed for each dish of embryos (see Methods).

TABLE 4

Treatment	24 hpa		48 hpa		
	Individual counts	Average	Individual counts	Average	
Control	81 131 88 113 184 136 95 171	125	68 49 45 37 65 43 10 108	53	
Concanamycin-exposed	79 26 45 53	51	39 58 51 46	49	

Description of Table 4:: Analysis of Medial Region Proliferation

[0527] Cells positive for the H3P proliferation marker were counted in control and concanamycin-exposed larvae in a

square region, with each side equal to the dorso-ventral height of the tail, located immediately posterior to the amputation plane.

[0528] In control amputated larvae, by 48 hpa, the number of proliferating cells in the mid-flank drops by a factor of 2.3. In contrast, in concanamycin-exposed larvae, the number does not change significantly between 24 and 48 hpa. It is important to note that there are two separate V-ATPase dependent events here. First, there is a general and immediate (by 24 hpa) 3-fold reduction in the number of proliferating cells after concanamycin exposure. This mild reduction of mitosis does not impact general development and is far smaller than the 10:1 reduction in proliferation in the regeneration bud. Thus, while the V-ATPase may be involved in proliferation in general, it is absolutely central to the up-regulation of proliferation in the bud.

[0529] Secondly, the normal down-regulation of proliferation at 48 hours does not occur when the V-ATPase is inhibited. This reduction is initiated by regeneration (is not local to the flank), and is V-ATPase-dependent. A model whereby the V-ATPase inhibition maintains proliferating cells locally in the flank is ruled out by the general negative effect of concanamycin on proliferation (opposite to the maintenance observed) and the fact that a local model would predict a further 3-fold reduction of the proliferating cell numbers in the flank, which is not observed. ANOVA analysis (N=24, F=11.02, p=0.0002) indicates a significant difference in the number of H3P-positive cells in the flank at 24 vs. 48 hours due to V-ATPase inhibition.

TABLE 5

Non-limiting examples of iontransporters are provided below. The nucleotide sequences, amino acid sequences, and other materials provided in these accession numbers are hereby incorporated by reference in their entirety

Ion Transporter	DDBJ/EMBL/GeneBank No. (nucleotide and amino acid sequences.)
TRPV1	DQ898279 (human); NM_001001445 (mouse)
CaV P/Q type (CACNA1A)	NM_000068 (human); NM_007578 (mouse)
Ca-ATPase (ATP2B1)	NM_001001323 (human); NM_001016839 (frog)
Na/Ca exchanger (SLC8A1)	NM_021097 (human); NM_011406 (mouse)

TABLE 5-continued

Non-limiting examples of iontransporters are provided below. The nucleotide sequences, amino acid sequences, and other materials provided in these accession numbers are hereby incorporated by reference in their entirety

Ion Transporter	DDBJ/EMBL/GeneBank No. (nucleotide and amino acid sequences.)
KCNQ1 (KvLQT1)	NM_000218 (human); NM_008434 (mouse)
KCNH2 (ERG)	NM_000238 (human); NM_013569 (mouse)
KCNK1	NM_002245 (human); NM_001011490 (frog)
KATP (Kir6.2)	NM_000891 (human); NM_008425 (mouse)
ROMK	NM_000220 (human); NM_019659 (mouse)
V-type H-ATPase	NM_005177, (for all subunits see Beyenbach et al., J Exp Biol. 2006 Feb; 209(Pt 4): 577-89, incorporated by reference herein)
Na/H exchanger (NHE-1)	S68616 (human); U51112 (mouse)
F-type H-ATPase (PMA1)	X03534 (<i>Saccharomyces cerevisiae</i>)
NaV (SCN5A)	NM_000335 (human); NM_021544 (mouse)
Na dependent anion exchanger (SLC9A1)	NM_003047 (human); NM_016981 (mouse)

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 52

<210> SEQ ID NO 1

<211> LENGTH: 2520

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

```

atgaagaaat ggagcagcac agacttgggg gcagctgcgg acccactccg aaaggacacc      60
tgcccagacc ccctggatgg agaccctaac tccaggccac ctccagccaa gccccagctc      120
tccacggcca agagccgcac cggctttt gggaaagggtg actcgagga ggcttcccg      180
gtggattgcc ctcacgagga aggtgagctg gactcctgcc cgaccatcac agtcagccct      240
gttatcacca tccagaggcc aggagacggc cccaccggtg ccaggctgt gtcccaggac      300
tctgtcgccg ccagcacgca gaagaccctc aggctctatg atcgcaggag tatctttgaa      360
gccgttgctc agaataactg ccaggatctg gagagccctgc tgctcttct gcagaagagc      420
aagaaggcacc tcacagacaa cgagttcaa gaccctgaga cagggaaagac ctgtctgtg      480
aaagccatgc tcaacctgca cgacggacag aacaccacca tccccctgt cctggagatc      540
gcgeggcaaa cggacagect gaaggagctt gtcaacgcca gctacacgga cagctactac      600
aaggggccaga cagcactgca catgcccattc gagagacgca acatggccct ggtgaccctc      660
ctgggtggaga acggagcaga cgtccaggct gggccccatg gggacttctt taagaaaacc      720
aaaggggccgc ctggattcta ctgcgtgaa ctgcccctgt ccctggccgc gtgcaccaac      780
cagctgggca tcgtgaagtt cctgctgcag aactcctggc agacggccga catcagcgcc      840
agggactcgg tggcaaacac ggtgctacac gcccctgggg aggtggccga caacacggcc      900
gacaacacga agtttgtgac gagcatgtac aatgagatc tgatcctggg ggccaaactg      960

```

-continued

cacccgacgc tgaagctgga ggagctcacc aacaagaagg gaatgacgcc gctggctctg	1020
gcagctggga cgggaaagat cggggcttgc gccttatattc tccagcggga gatccaggag	1080
cccgagtgc ggcacctgtc caggaagttc accgagtgcc cctgcggcc cgtgcactcc	1140
tgcgtgtacg acctgtcctg catgcacacc tgcgagaaga actcggtgt ggagggatc	1200
gcctacagca gcagcgagac ccctaattgc cagcacatgc tcttgggta gccgctgaac	1260
cgactcctgc aggacaagtgg gacagatcc gtcaagcgca tcttctactt caacttcctg	1320
gtctactgcc tgtacatgtat catcttcacc atggctgcct actacaggcc cgtggatggc	1380
ttgcctccct ttaagatggaaa aaaaactggaa gactattcc gagttactgg agagatcctg	1440
tctgtgttag gaggagtcta cttcttttc cgagggattc agtatttctt gcagaggccg	1500
ccgtcgatga agaccctgtt tgtgaacagc tacagtggaa tgctttctt tctgcgtca	1560
ctgttcatgc tggccacccgt ggtgctgtac ttcaaggccacc tcaaggagta tgtggcttcc	1620
atggtattctt ccctggcattt gggctggacc aacatgtctt actacaccgg cgggtttccag	1680
cagatggca tctatgcgtt catgatagaa aagatgtacc tgagagacccgt gtgcgttcc	1740
atgtttgtctt acatcgctt cttgttcggg tttccacag cgggtggatc gctgattgaa	1800
gacgggaaga atgactccctt gcccgtctgag tccacgtcgc acagggtggcg gggccctgcc	1860
tgcaggcccccc ccatggatc ctacaacaggc ctgtactcca cctgcctggaa gctgttcaag	1920
ttcaccatcg gcatggcga cctggagttc actgagaact atgacttcaa ggctgttcc	1980
atcatccatgc tgcgtggctgtt tgcatttccatcc tccatgtccaa catgtccatc	2040
gccctcatgg gtgagactgtt caacaagatc gacaggaga gcaaggacat ctggagactg	2100
cagagagcca tcaccatctt ggacacgggg aagagcttcc ttaagtgcattt gaggaaggcc	2160
ttccgcgtcg gcaagctgtt gcaagggtggg tacacaccgtt atggcaagggcc gactaccgg	2220
tggtgcttca ggggtggcga ggtgaactgg accacccgtt gaccaacatggcc gggcatcatc	2280
aacgaagacc cgggcaactg tgaggcgttc aagcgcaccc tgagcttctt cctgeggatc	2340
agcagagttt caggcagaca ctggaaagac tttggccctgg tccccctttt aagagaggca	2400
agtgcgtcgatc ataggcgttc tgctcagccc gaggaagttt atctgegaca gttttaggg	2460
tctctgaagc cagaggacgc tgaggcttc aagagtccgtt ccgcgttccgg ggagaagtgaa	2520

<210> SEQ_ID NO 2
<211> LENGTH: 839
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

Met Lys Lys Trp Ser Ser Thr Asp Leu Gly Ala Ala Ala Asp Pro Leu

1 5 10 15

Arg Lys Asp Thr Cys Pro Asp Pro Leu Asp Gly Asp Pro Asn Ser Arg

20 25 30

Pro Pro Pro Ala Lys Pro Gln Leu Ser Thr Ala Lys Ser Arg Thr Arg

35 40 45

Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Phe Pro Val Asp Cys Pro

50 55 60

His Glu Glu Gly Glu Leu Asp Ser Cys Pro Thr Ile Thr Val Ser Pro

65 70 75 80

-continued

Val Ile Thr Ile Gln Arg Pro Gly Asp Gly Pro Thr Gly Ala Arg Leu			
85	90	95	
Leu Ser Gln Asp Ser Val Ala Ala Ser Thr Glu Lys Thr Leu Arg Leu			
100	105	110	
Tyr Asp Arg Arg Ser Ile Phe Glu Ala Val Ala Gln Asn Asn Cys Gln			
115	120	125	
Asp Leu Glu Ser Leu Leu Phe Leu Gln Lys Ser Lys Lys His Leu			
130	135	140	
Thr Asp Asn Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu			
145	150	155	160
Lys Ala Met Leu Asn Leu His Asp Gly Gln Asn Thr Thr Ile Pro Leu			
165	170	175	
Leu Leu Glu Ile Ala Arg Gln Thr Asp Ser Leu Lys Glu Leu Val Asn			
180	185	190	
Ala Ser Tyr Thr Asp Ser Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile			
195	200	205	
Ala Ile Glu Arg Arg Asn Met Ala Leu Val Thr Leu Leu Val Glu Asn			
210	215	220	
Gly Ala Asp Val Gln Ala Ala Ala His Gly Asp Phe Phe Lys Lys Thr			
225	230	235	240
Lys Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala			
245	250	255	
Ala Cys Thr Asn Gln Leu Gly Ile Val Lys Phe Leu Leu Gln Asn Ser			
260	265	270	
Trp Gln Thr Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val			
275	280	285	
Leu His Ala Leu Val Glu Val Ala Asp Asn Thr Ala Asp Asn Thr Lys			
290	295	300	
Phe Val Thr Ser Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu			
305	310	315	320
His Pro Thr Leu Lys Leu Glu Leu Thr Asn Lys Lys Gly Met Thr			
325	330	335	
Pro Leu Ala Leu Ala Ala Gly Thr Gly Lys Ile Gly Val Leu Ala Tyr			
340	345	350	
Ile Leu Gln Arg Glu Ile Gln Glu Pro Glu Cys Arg His Leu Ser Arg			
355	360	365	
Lys Phe Thr Glu Trp Ala Cys Gly Pro Val His Ser Ser Leu Tyr Asp			
370	375	380	
Leu Ser Cys Ile Asp Thr Cys Glu Lys Asn Ser Val Leu Glu Val Ile			
385	390	395	400
Ala Tyr Ser Ser Ser Glu Thr Pro Asn Arg His Asp Met Leu Leu Val			
405	410	415	
Glu Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp Arg Phe Val Lys			
420	425	430	
Arg Ile Phe Tyr Phe Asn Phe Leu Val Tyr Cys Leu Tyr Met Ile Ile			
435	440	445	
Phe Thr Met Ala Ala Tyr Tyr Arg Pro Val Asp Gly Leu Pro Pro Phe			
450	455	460	
Lys Met Glu Lys Thr Gly Asp Tyr Phe Arg Val Thr Gly Glu Ile Leu			
465	470	475	480
Ser Val Leu Gly Gly Val Tyr Phe Phe Arg Gly Ile Gln Tyr Phe			

-continued

485	490	495
Leu Gln Arg Arg Pro Ser Met Lys Thr Leu Phe Val Asn Ser Tyr Ser		
500	505	510
Glu Met Leu Phe Phe Leu Gln Ser Leu Phe Met Leu Ala Thr Val Val		
515	520	525
Leu Tyr Phe Ser His Leu Lys Glu Tyr Val Ala Ser Met Val Phe Ser		
530	535	540
Leu Ala Leu Gly Trp Thr Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln		
545	550	555
560		
Gln Met Gly Ile Tyr Ala Val Met Ile Glu Lys Met Ile Leu Arg Asp		
565	570	575
Leu Cys Arg Phe Met Phe Val Tyr Ile Val Phe Leu Phe Gly Phe Ser		
580	585	590
Thr Ala Val Val Thr Leu Ile Glu Asp Gly Lys Asn Asp Ser Leu Pro		
595	600	605
Ser Glu Ser Thr Ser His Arg Trp Arg Gly Pro Ala Cys Arg Pro Pro		
610	615	620
Asp Ser Ser Tyr Asn Ser Leu Tyr Ser Thr Cys Leu Glu Leu Phe Lys		
625	630	635
640		
Phe Thr Ile Gly Met Gly Asp Leu Glu Phe Thr Glu Asn Tyr Asp Phe		
645	650	655
Lys Ala Val Phe Ile Ile Leu Leu Ala Cys Val Ile Leu Thr Tyr		
660	665	670
Ile Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val Asn		
675	680	685
Lys Ile Ala Gln Glu Ser Lys Asp Ile Trp Lys Leu Gln Arg Ala Ile		
690	695	700
Thr Ile Leu Asp Thr Glu Lys Ser Phe Leu Lys Cys Met Arg Lys Ala		
705	710	715
720		
Phe Arg Ser Gly Lys Leu Leu Gln Val Gly Tyr Thr Pro Asp Gly Lys		
725	730	735
Asp Asp Tyr Arg Trp Cys Phe Arg Val Asp Glu Val Asn Trp Thr Thr		
740	745	750
Trp Asn Thr Asn Val Gly Ile Ile Asn Glu Asp Pro Gly Asn Cys Glu		
755	760	765
Gly Val Lys Arg Thr Leu Ser Phe Ser Leu Arg Ser Ser Arg Val Ser		
770	775	780
Gly Arg His Trp Lys Asn Phe Ala Leu Val Pro Leu Leu Arg Glu Ala		
785	790	795
800		
Ser Ala Arg Asp Arg Gln Ser Ala Gln Pro Glu Glu Val Tyr Leu Arg		
805	810	815
Gln Phe Ser Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Ser		
820	825	830
Pro Ala Ala Ser Gly Glu Lys		
835		

```

<210> SEQ_ID NO 3
<211> LENGTH: 2520
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 3

```

-continued

atggagaaat gggctagett agactcgat gaatctgagc ccccagccca agagaactcc	60
tgcggggacc ctccagacag agaccctaac tccaagccgc ctccagccaa gccccacatc	120
tttgctacca ggagtgcac ccggctttt gggagggtg actcagaaga ggcctctccc	180
atggactgcc cttatgagga aggccggctg gcctcctgcc ctatcatcac cgtcagctct	240
gttgtcactc tccagaggc tggatgga cctacctgtc tcaggcagac atcccaggac	300
tctgtctcca ctgggttga gacgccccca aggctctatg atcgcaggag catcttcgac	360
gctgtggctc agagcaactg ccaggagctg gagagcctgc tgtccttcgc gcagaagagc	420
aagaagcgc aagaactgacag cgagttcaaa gaccagaga cgggaaagac ctgtctgctc	480
aaagccatgc tcaatctgca caatggcag aacgacacca ttgctctgct cctggacatt	540
gccggagaaga cagatagect gaaggagttt gtcaatgcca gtcacacaga cagctactac	600
aaggggccaga cagcattaca cattgccatt gaaaggcggaa acatggact ggtgaccctc	660
ttgggtggaga atggagcaga tgtccaggct gtcgttaacg gggacttctt caagaaaacc	720
aaaggaggc ctggcttcta ctttggtagt ctgccccgt ccctggctgc gtgcaccaac	780
cagctggca ttgtgaagtt cctgctgcag aactcctggc agcctgcaga catcagtgc	840
cgggattcgg tggcaacac ggtgctgcac gcccttggg aggtggcaga taacacagct	900
gacaacacca agtcgtgac aaacatgtac aacgagatcc tgatcctggg ggc当地actc	960
cacccacac tgaagctaga agaactcacc aacaagaagg ggcttacacc gctggctctg	1020
gctgccagca gtggaaagat tgggtcttg gcctacatc tccagaggaa gatccacgaa	1080
ccagagtgcc ggcacactgca caggaagttt actgaatggg cctatggcc cgtgcactcc	1140
tcccttatg acctgtctgc cattgacacc tgtgagaaga attcagtgt ggagggtgate	1200
gcctacagta gcagtggagac ccccaaccgc cacgacatgc ttctcgatgg gccc当地gaac	1260
cgactcctgc aggacaagtg ggacagattt gtcaagcgc tcttctactt caacttcttc	1320
gtctactgtctgtatgat catcttcacc acggctgcctt actatcgcc tgtggaaaggc	1380
ttgccccct ataagctgaa taacaccgtt gggacttatt tccgtgtac tggagagatc	1440
ctgtctgtgtt caggaggatg ctacttcttc ttccggggaa tccagttttt cctgcagagg	1500
cgaccatccc tcaagagttt gtttggac agtacagtg agataactttt ctttgcacag	1560
tcactgttca tgctgggtgc tgggtactg tacttcagcc atcgcaagga gtatgtggct	1620
tccatgggtgt tctccctggc catggctgg accaacatgc tctactacac ccgaggattc	1680
cagcagatgg gcatctatgc tgcgtatgat gagaagatga tccctcagaga cctgtgtcgg	1740
tttatgttcg tctacctcg tttttttttt ggatttccca cagccgtagt gacactgtatc	1800
gaggatgggaa agaataactc actgcctgtg gatgtccac cacacaagtg tcggggatct	1860
gcctgcaggc caggttaactc ttacaacacgc ctgtattcca catgtctgg gctgttcaag	1920
ttcaccatcg gcatgggtga cctggagttt accgagaact atgacttcaa ggctgttctc	1980
atcatcctgt tactggccta tggatttctc acctacatcc tccctgtccaa catgttcatc	2040
gctctcatgg gcgagactgt caacaagattt gcacaagaga gcaagaacat ctggaaagct	2100
cagcggccca tcaaccatctt ggatacagag aagagttcc tgaagtgcat gaggaaggcc	2160
tcccgctccg gcaagctgtt gcaagggtggg ttcacgccc acggcaagga tgacttccgg	2220
tggtgcttca ggggtggatga ggtgaactgg actacctgg aacccaacgt gggcatcatc	2280

-continued

```

aacgaggacc caggcaactg tgagggcgtc aagcgcaccc tgagcttctc cctgccccgtca 2340
ggcccgatgttt cagggagaaa ctggaaagaac ttggccctgg ttccccccttctt gaggggacgca 2400
agcactcgag ataggcatag caccggccg gaagaagtgc agctgaagca ctatacggg 2460
tcccttaagc cagaggatgc tgaggtcttc aaggattcca tggccccagg ggagaaatga 2520

<210> SEQ ID NO 4
<211> LENGTH: 839
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 4

Met Glu Lys Trp Ala Ser Leu Asp Ser Asp Glu Ser Glu Pro Pro Ala
1 5 10 15

Gln Glu Asn Ser Cys Pro Asp Pro Pro Asp Arg Asp Pro Asn Ser Lys
20 25 30

Pro Pro Pro Ala Lys Pro His Ile Phe Ala Thr Arg Ser Arg Thr Arg
35 40 45

Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Ser Pro Met Asp Cys Pro
50 55 60

Tyr Glu Glu Gly Gly Leu Ala Ser Cys Pro Ile Ile Thr Val Ser Ser
65 70 75 80

Val Val Thr Leu Gln Arg Ser Val Asp Gly Pro Thr Cys Leu Arg Gln
85 90 95

Thr Ser Gln Asp Ser Val Ser Thr Gly Val Glu Thr Pro Pro Arg Leu
100 105 110

Tyr Asp Arg Arg Ser Ile Phe Asp Ala Val Ala Gln Ser Asn Cys Gln
115 120 125

Glu Leu Glu Ser Leu Leu Ser Phe Leu Gln Lys Ser Lys Lys Arg Leu
130 135 140

Thr Asp Ser Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu
145 150 155 160

Lys Ala Met Leu Asn Leu His Asn Gly Gln Asn Asp Thr Ile Ala Leu
165 170 175

Leu Leu Asp Ile Ala Arg Lys Thr Asp Ser Leu Lys Gln Phe Val Asn
180 185 190

Ala Ser Tyr Thr Asp Ser Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile
195 200 205

Ala Ile Glu Arg Arg Asn Met Ala Leu Val Thr Leu Leu Val Glu Asn
210 215 220

Gly Ala Asp Val Gln Ala Ala Asn Gly Asp Phe Phe Lys Lys Thr
225 230 235 240

Lys Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala
245 250 255

Ala Cys Thr Asn Gln Leu Ala Ile Val Lys Phe Leu Leu Gln Asn Ser
260 265 270

Trp Gln Pro Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val
275 280 285

Leu His Ala Leu Val Glu Val Ala Asp Asn Thr Ala Asp Asn Thr Lys
290 295 300

Phe Val Thr Asn Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu
305 310 315 320

```

-continued

His	Pro	Thr	Leu	Lys	Leu	Glu	Glu	Leu	Thr	Asn	Lys	Lys	Gly	Leu	Thr
325															335
Pro	Leu	Ala	Leu	Ala	Ala	Ser	Ser	Gly	Lys	Ile	Gly	Val	Leu	Ala	Tyr
340															350
Ile	Leu	Gln	Arg	Glu	Ile	His	Glu	Pro	Glu	Cys	Arg	His	Leu	Ser	Arg
355															365
Lys	Phe	Thr	Glu	Trp	Ala	Tyr	Gly	Pro	Val	His	Ser	Ser	Leu	Tyr	Asp
370															380
Leu	Ser	Cys	Ile	Asp	Thr	Cys	Glu	Lys	Asn	Ser	Val	Leu	Glu	Val	Ile
385															400
Ala	Tyr	Ser	Ser	Glu	Thr	Pro	Asn	Arg	His	Asp	Met	Leu	Leu	Val	
405															415
Glu	Pro	Leu	Asn	Arg	Leu	Leu	Gln	Asp	Lys	Trp	Asp	Arg	Phe	Val	Lys
420															430
Arg	Ile	Phe	Tyr	Phe	Asn	Phe	Phe	Val	Tyr	Cys	Leu	Tyr	Met	Ile	Ile
435															445
Phe	Thr	Thr	Ala	Ala	Tyr	Tyr	Arg	Pro	Val	Glu	Gly	Leu	Pro	Pro	Tyr
450															460
Lys	Leu	Asn	Asn	Thr	Val	Gly	Asp	Tyr	Phe	Arg	Val	Thr	Gly	Glu	Ile
465															480
Leu	Ser	Val	Ser	Gly	Gly	Val	Tyr	Phe	Phe	Phe	Arg	Gly	Ile	Gln	Tyr
485															495
Phe	Leu	Gln	Arg	Arg	Pro	Ser	Leu	Lys	Ser	Leu	Phe	Val	Asp	Ser	Tyr
500															510
Ser	Glu	Ile	Leu	Phe	Phe	Val	Gln	Ser	Leu	Phe	Met	Leu	Val	Ser	Val
515															525
Val	Leu	Tyr	Phe	Ser	His	Arg	Lys	Glu	Tyr	Val	Ala	Ser	Met	Val	Phe
530															540
Ser	Leu	Ala	Met	Gly	Trp	Thr	Asn	Met	Leu	Tyr	Tyr	Thr	Arg	Gly	Phe
545															560
Gln	Gln	Met	Gly	Ile	Tyr	Ala	Val	Met	Ile	Glu	Lys	Met	Ile	Leu	Arg
565															575
Asp	Leu	Cys	Arg	Phe	Met	Phe	Val	Tyr	Leu	Val	Phe	Leu	Phe	Gly	Phe
580															590
Ser	Thr	Ala	Val	Val	Thr	Leu	Ile	Glu	Asp	Gly	Lys	Asn	Asn	Ser	Leu
595															605
Pro	Val	Glu	Ser	Pro	Pro	His	Lys	Cys	Arg	Gly	Ser	Ala	Cys	Arg	Pro
610															620
Gly	Asn	Ser	Tyr	Asn	Ser	Leu	Tyr	Ser	Thr	Cys	Leu	Glu	Leu	Phe	Lys
625															640
Phe	Thr	Ile	Gly	Met	Gly	Asp	Leu	Glu	Phe	Thr	Glu	Asn	Tyr	Asp	Phe
645															655
Lys	Ala	Val	Phe	Ile	Ile	Leu	Leu	Leu	Ala	Tyr	Val	Ile	Leu	Thr	Tyr
660															670
Ile	Leu	Leu	Leu	Asn	Met	Leu	Ile	Ala	Leu	Met	Gly	Glu	Thr	Val	Asn
675															685
Lys	Ile	Ala	Gln	Glu	Ser	Lys	Asn	Ile	Trp	Lys	Leu	Gln	Arg	Ala	Ile
690															700
Thr	Ile	Leu	Asp	Thr	Glu	Lys	Ser	Phe	Leu	Lys	Cys	Met	Arg	Lys	Ala
705															720
Phe	Arg	Ser	Gly	Lys	Leu	Leu	Gln	Val	Gly	Phe	Thr	Pro	Asp	Gly	Lys

-continued

725	730	735	
Asp Asp Phe Arg Trp Cys Phe Arg Val Asp Glu Val Asn Trp Thr Thr			
740	745	750	
Trp Asn Thr Asn Val Gly Ile Ile Asn Glu Asp Pro Gly Asn Cys Glu			
755	760	765	
Gly Val Lys Arg Thr Leu Ser Phe Ser Leu Arg Ser Gly Arg Val Ser			
770	775	780	
Gly Arg Asn Trp Lys Asn Phe Ala Leu Val Pro Leu Leu Arg Asp Ala			
785	790	795	800
Ser Thr Arg Asp Arg His Ser Thr Gln Pro Glu Glu Val Gln Leu Lys			
805	810	815	
His Tyr Thr Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Asp			
820	825	830	
Ser Met Ala Pro Gly Glu Lys			
835			

<210> SEQ ID NO 5

<211> LENGTH: 7807

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

cgcacccctcc	ttcogccccct	ctttctccggg	ggtcagccag	gaagatgtcc	cgagctgcta	60
tccccggctc	ggcccgggca	gcgcgccttc	gagcccccca	cccgaggcgc	cgagccggcg	120
ccgccccatg	ggctggggcg	tggagcgtct	ccgcagtcgt	agctccagcc	gccgegcgtcc	180
cagccccggc	agccctcagca	tca g cggcgg	cg g cggcgggc	ggcggcggcgc	tcttcogcat	240
cg ttc gcccgc	agegtaaacc	ggagccctt	gtctttgc	aatggcccg	cttcggagac	300
gagatgcggg	cccgctacgg	gggaggaggc	tccggggcag	ccgcgggggt	gttcgtggc	360
agcggaggcg	ggcggaggcgc	cgggggcagc	cg g caggcgc	ggcagecccg	ggcgc aaagg	420
atgtacaagc	agtcaatggc	gcagagagcg	cg g accatgg	cactctacaa	ccccatcccc	480
gtccgacaga	actgcctc ac	ggtaaccgg	tctctttcc	tcttcagcga	agacaacgtg	540
gtgagaaaat	acgc caaaaa	gatcacgaa	tggcctccct	ttgaatata	tttttagcc	600
accatcatag	cgaattgcat	cgtccctcgca	ctggagcagc	atctgcctga	t gatgacaag	660
accccgatgt	ctgaacggct	ggatgacaca	gaaccata	tca ttggaa	ttttttttc	720
gaggctggaa	t taaaatcat	tgc ct tggg	tttgccttcc	aca aaggc	ctacttgg	780
aatggctgga	atgtcatgga	ctttgttgt	gtgctaaccg	gcatcttgc	gacagtgg	840
acggagttt	ac t tacggac	gtc gagggc	gtt c g a g t	tc gggccgc	ca a getgg	900
tctggaatcc	ca agtttaca	atgcgtctg	a a g t c g a t c a	t g a a g g c	gatcccttgc	960
ctgc a gatcg	gcctccct	at ttttgc	a tc cttat	t tgcaatcat	agg gtt	1020
ttttatatgg	gaaaatttca	taccac	tgc	tttgaagagg	ggacagatg	1080
gagtctccgg	c tccatgtgg	gacagaagag	cccgccccca	cctgccccaa	tgggaccaaa	1140
tgtcagccct	actgggaagg	gccc aacaac	ggatcactc	agttc gacaa	c a t c c t g t t	1200
gc a gt gctg a	ctgtttcca	gtcataacc	at g g a a g g	gtactgatct	cctctacaat	1260
agcaacgatg	cctcaggaa	cacttggaa	tgg t g t a c t	t c a t c c c t	catcatcatc	1320
ggctcccttt	t t a t g c t g a a	cctt g t g c t g	ggt g t g c t g t	c a g g g g a g t t	t g c c a a a g a a	1380

-continued

agggaaacggg tggagaaccg gcgggcgttt ctgaagctga ggccgcacaaca acagattgaa	1440
cgtgagctca atgggtacat ggaatggatc tcaaaaagcag aagagggtat cctcgccgag	1500
gatgaaactg acggggagca gaggcatccc tttatggag ctctgcggag aaccaccata	1560
aagaaaagca agacagattt gctcaacccc gaagaggctg aggtcagct ggctgatata	1620
gcctctgtgg gttctccctt cgcccgagcc agcattaaaa gtgcgaagct ggagaactcg	1680
accttttttc acaaaaagga gaggaggatg cgtttctaca tccgcgcgt ggtcaaaaact	1740
caggccttct actggactgt actcagttt gtagctctca acacgctgtg tggtgttatt	1800
gttcaactaca accagccccga gtggctctcc gacttccttt actatgcaga attcatttc	1860
ttaggactct ttatgtccga aatgtttata aaaatgtacg ggcttggac gcggccttac	1920
ttccactctt cttcaactg ctttgcgtt ggggttatca ttgggagcat ctgcggaggtc	1980
atctgggctg tcataaaaacc tggcacatcc tttggaatca gcgtgttacg agccctcagg	2040
ttattgcgtta tttcaaaagt cacaaggatc tggcgcatttc tcagaaacct ggtcgctct	2100
ctccctcaact ccatgaagtc catcatcage ctgttgtttc tcctttctt gttcattgtc	2160
gtcttcgccc ttttggaaat gcaacttttc ggccggccagt ttaatttgcg tgaaggact	2220
cctcccacca acttcgatac tttccagca gcaataatga cggtgtttca gatcctgacg	2280
ggcgaagact ggaacggagt catgtacgac gggatcaagt ctcagggggg cgtgcaggc	2340
ggcatggtgt tctccatcta tttcattgtc ctgacgctct ttgggaaacta caccctcctg	2400
aatgtgttct tggccatcgc tgggacaaat ctggccaaacg cccaggagct caccaaggac	2460
gagcaagagg aagaagaagc agcgaaccag aaacttgcac tacagaaagc caaggagggt	2520
gcagaagtga gtccctgtc cgccggccaaat atgtctatag ctgtgaaaga gcaacagaag	2580
aatcaaaaagc cagccaaatc cgtgtggag cagccggacca gtgagatgcg aaagcagaac	2640
ttgtgtggca gccggggaggc cctgtataac gaaatggacc cggacgacgc ctggaggct	2700
gcctacacgc ggcacctgcg gccagacatg aagacgact tggacggcc gctgggttg	2760
gaccggcagg agaaccgcaaa caacaacacc aacaagagcc gggccggccga gcccacccgt	2820
gaccaggcgc tcggccagca gggccggcggag gacttcctca ggaaacaggc ccgctaccac	2880
gatcgccccc gggaccccg cggctcgccg ggcctggacg cacggaggcc ctggggggga	2940
agccaggaggcc cggagctgag cggggaggga ccctacggcc gcgagtcgga ccaccacgcc	3000
cgccggggca gcctgggacca accggggttc tggggggcg aggccgacgc aggcaaggcc	3060
ggggacccccc accggaggca cgtgcacccg cagggggca gcaaggaggag ccgcacccgg	3120
tcccccgcga cggggccggc cggggaggat cgacgtcatc ggcgcacccg cagcccccgg	3180
gaggagggtc cggaggacaa ggccggagcgg agggcggggc accgcgaggc cagccggccg	3240
gccccggggc gcgagggega gggcgaggccc cccgacgggg gcgagcgcg gagaaggcac	3300
cgccatggcg cttccagccac gtacgggggg gacgcggggc gggaggacaa ggacggagg	3360
catcgaggaa ggaaaggagaa ccagggttcc ggggtccctg tgcggggccca caacctgtca	3420
accacccggc caatccagca ggacctgggc cgccaaagacc cacccctggc agaggatatt	3480
gacaacatga agaacaacaa gctggccacc gggaggatcgcc cggctccca cggcagcctt	3540
ggccacccgcg gcctgccccca gagcccgacca aagatggaa acagcaccga ccccgcccc	3600
atgctggcca tccctgccccat ggccaccaac ccccaagaacg cggccagccg cggacgcccc	3660

-continued

aacaaccgcg ggaaccatc caatccggc cccccaaga cccccagaa tagccttac 3720
gtcaccacc accccggcac ccagaccaat tcagctaaga ctgccaggaa acccgaccac 3780
accacagtgg acatcccccc agcctgcca ccccccctca accacaccgt cgtacaagtg 3840
aacaaaaacg ccaaccaga cccactgcc aaaaaagagg aagagaagaa ggaggaggag 3900
gaagacgacc gtgggaaaga cgccctaaag ccaatgcctc cctatactc catttcatc 3960
ctgtccacg ccaacccttc tcgcccctg tgccattaca tcctgaacct gcgcacttt 4020
gagatgtca tcctcatggt cattgccatg agcagcatcg ccctggccgc cgaggaccct 4080
gtgcagccca acgcaccccg gaacaacgtg ctgcgatact ttgactacgt ttttacaggc 4140
gtcttaccc ttgagatggt gatcaagatg attgacccgg ggctcgctt gcatagggt 4200
gcctacttcc gtgacccctcg gaatattctc gacttcatag tggctagtgg ggccctggta 4260
gccttgcct tcactggcaa tagcaaagga aaagacatca acacgattaa atccctccga 4320
gtcctccggg tgctacgacc tctaaaacc atcaagggc tgccaaagct caaggctgtg 4380
tttgactgtg tggtaactc actaaaaaac gtctcaaca tcctcatcgt ctacatgcta 4440
ttcatgttca tcttcggcgt ggtggctgtg cagctttca agggaaaatt ctccactgc 4500
actgacgagt ccaaagagtt tgagaaagat tgtcgaggca aataccctct ctacgagaag 4560
aatgaggatgta aggcgcgaga ccgggagtt aagaagtatg aattccattt cgacaatgt 4620
ctgtgggctc tgctgaccct cttcacccgt tccacgggag aaggctggcc acaggctcc 4680
aagcattccgg tggacgcccac ctttgagaac cagggcccca gccccggta ccgcattggag 4740
atgtccattt tctacgtcgt ctactttgtg gtgttccct tcttcttgc caatatcttt 4800
gtggccttga tcatcatcac cttccaggag caaggggaca agatgtatg ggaatacagc 4860
ctggagaaaa atgagaggc ctgcattgtat ttcgcattca ggcggccaaagcc gctgaccgc 4920
cacatgcgc agaacaagca gagttccag taccgcattt ggcagttcg ggtgtctcc 4980
ccttcgagt acacgatcat ggccatgatc gcctcaaca ccattgtct tatgtatgaa 5040
ttctatgggg cttctgttgc ttatgaaaat gcctgggg tggtaacat cgtcttacc 5100
tccctttct ctctggaaat tggctgtggaa gtcattggc ttggattctt gaattatttc 5160
cgcgatccctt ggaacatctt cgactttgtg actgttctgg gcagcatcac cgatattcc 5220
gtgactgagt ttgggataaa cttcatcaac ctgagtttc tccgccttcc cggagctgc 5280
cggttcatca aacttcccg tcagggttac accatccgc ttcttcttgc gaccttgc 5340
cagtcattca aggccctgcc ttatgtctgt ctgctgatcg ccattgtctt cttcatctat 5400
gcattcatggt ggtatgcagggt ttggtaac attggcatcg acgtggagga cgaggacagt 5460
gatgaagatg agttccaaat cactgagcac aataacttcc ggaccttctt ccaggccctc 5520
atgttctcttcc tccggatgtc caccggggaa gcttggcaca acatcatgtt ttcctgcctc 5580
agoggggaaac cgtgtgataa gaactctggc atctgactc gagagtgtgg caatgaattt 5640
gttattttt actttgttcc cttcatcttcc ctgtgtcgat ttctgtatgtt gaatctttt 5700
gtggccgtca tcattggacaa ctttgatgc ctcaccccgag actctccat cttggggcc 5760
caccacccggg atgagatcgatc gctgtgttgc gccgagttatg accccgcgc ttgegggtcg 5820
attcattata aggatatgtt cagtttattt cagtaatat ctccccctctt cgggttaggc 5880
aagaaatgtc ctcatacggtt tgcttgcag cggcttgc ggtatggaccc gcccgtgc 5940

-continued

gatgacaaca ccgtccactt caattccacc ctcatggctc tcatccgcac agccctggac	6000
atcaagattt ccaaggagg agccgacaaa cagcagatgg acgctgagct gcggaaggag	6060
atgatggcga tttggccaa tctgtcccag aagacgctag acctgcttgt cacacctac	6120
aagtccacgg acctcaccgt ggggaagatc tacgcagcca tcatgtatcat ggagtactac	6180
cggcagagca aggccaagaa gctgcaggcc atgcgcgagg agcaggaccg gacaccctc	6240
atgttccagc gcatggagcc cccgtccccca acgcaggaag ggggacattgg ccagaacgcc	6300
ctccctcca cccagctgga cccaggagga gcccgtatgg ctcacgaaag cggcctcaag	6360
gagagccctgt cctgggtgac ccagcgtgcc caggagatgt tccagaagac gggcacatgg	6420
agtccggAAC aaggcccccc taccgacatg cccaaacagcc agcctaactc tcagtcgt	6480
gagatgcgag agatgggcag agatggctac tccgacagcg agcactaccc ccccatggaa	6540
ggccaggggcc gggctgcctc catgccccgc ctccctgcag agaaccagag gagaaggggc	6600
cggccacgtg ggaataacct cagtaccatc tcagacacca gccccatgaa gcgttcagcc	6660
tccgtgtggc gccccaaaggc ccgacgcctg gacgattact cgctggagcg ggtcccgccc	6720
gaggagaacc agggggccca ccagcggccg cgcgaccgc gccaccgcgc ctctgagcgc	6780
tccctgggac gctacaccga tgtggacaca ggcttgggaa cagacctgag catgaccacc	6840
caatccgggg acctgcgcgc gaaggagcgg gaccaggagc gggggccggcc caaggatcgg	6900
aagcatcgac agcaccacca ccaccaccc caccacccac atccccggcc ccccgacaag	6960
gaccgctatg cccaggaacg gccggaccac gggccgggcac gggctcgggg ccagegctgg	7020
tcccgctcgc ccagcgaggg ccgagagcac atggcgccac ggcagtagtt ccgttaagtgg	7080
aageccagec ccctcaacat ctgggtaccag cactccgcgg cggggccgc gccagotccc	7140
ccagacccccc tccacccccc ggocacacgt gtcttatccc cctgtgtatcc gtaaggccgg	7200
cggctcgggg ccccccgcgc agcagcgcac gcagcagcagc cagcagcagg cggtgccag	7260
gccccggccgg gccccccacca gggccctcg gaggtaccca gggccacgg ccgagctct	7320
ggccggagat cggccgcaca cggggggccca cagcagcggc cgctcgccca ggatggagag	7380
gccccgtccca ggeccggccc ggagcgagtc ccccaaggcc tgctgacacg gccccggcc	7440
gtggccggca tctggccgc acgtgtccga gggggccccc ggtccccggc accatggcta	7500
ctaccggggc tccgactacg acgaggccga tggccgggc agcggggccg gcgaggaggc	7560
catggccggg gcctacgcacg cgccacccccc cgtacgacac gctcctcg ggcaccacgg	7620
gcgcctcgccc aggactcccc gggcctcggt cccggccctgc gcctcgccctt ctggcagcc	7680
cggcgactc cccaaacggct actacccggc gcacggactg gccaggcccc gccccgggg	7740
ctccaggaag ggcctgcacg aaccctacag cgagagtgac gatgatttgt gctaagcccg	7800
ggcgagg	7807

<210> SEQ ID NO 6
<211> LENGTH: 2261
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

Met	Ala	Arg	Phe	Gly	Asp	Glu	Met	Pro	Ala	Arg	Tyr	Gly	Gly	Gly
1					5			10				15		

-continued

Ser Gly Ala Ala Ala Gly Val Val Val Gly Ser Gly Gly Gly Arg Gly
20 25 30

Ala Gly Gly Ser Arg Gln Gly Gly Gln Pro Gly Ala Gln Arg Met Tyr
35 40 45

Lys Gln Ser Met Ala Gln Arg Ala Arg Thr Met Ala Leu Tyr Asn Pro
50 55 60

Ile Pro Val Arg Gln Asn Cys Leu Thr Val Asn Arg Ser Leu Phe Leu
65 70 75 80

Phe Ser Glu Asp Asn Val Val Arg Lys Tyr Ala Lys Lys Ile Thr Glu
85 90 95

Trp Pro Pro Phe Glu Tyr Met Ile Leu Ala Thr Ile Ile Ala Asn Cys
100 105 110

Ile Val Leu Ala Leu Glu Gln His Leu Pro Asp Asp Asp Lys Thr Pro
115 120 125

Met Ser Glu Arg Leu Asp Asp Thr Glu Pro Tyr Phe Ile Gly Ile Phe
130 135 140

Cys Phe Glu Ala Gly Ile Lys Ile Ile Ala Leu Gly Phe Ala Phe His
145 150 155 160

Lys Gly Ser Tyr Leu Arg Asn Gly Trp Asn Val Met Asp Phe Val Val
165 170 175

Val Leu Thr Gly Ile Leu Ala Thr Val Gly Thr Glu Phe Asp Leu Arg
180 185 190

Thr Leu Arg Ala Val Arg Val Leu Arg Pro Leu Lys Leu Val Ser Gly
195 200 205

Ile Pro Ser Leu Gln Val Val Leu Lys Ser Ile Met Lys Ala Met Ile
210 215 220

Pro Leu Leu Gln Ile Gly Leu Leu Leu Phe Phe Ala Ile Leu Ile Phe
225 230 235 240

Ala Ile Ile Gly Leu Glu Phe Tyr Met Gly Lys Phe His Thr Thr Cys
245 250 255

Phe Glu Glu Gly Thr Asp Asp Ile Gln Gly Glu Ser Pro Ala Pro Cys
260 265 270

Gly Thr Glu Glu Pro Ala Arg Thr Cys Pro Asn Gly Thr Lys Cys Gln
275 280 285

Pro Tyr Trp Glu Gly Pro Asn Asn Gly Ile Thr Gln Phe Asp Asn Ile
290 295 300

Leu Phe Ala Val Leu Thr Val Phe Gln Cys Ile Thr Met Glu Gly Trp
305 310 315 320

Thr Asp Leu Leu Tyr Asn Ser Asn Asp Ala Ser Gly Asn Thr Trp Asn
325 330 335

Trp Leu Tyr Phe Ile Pro Leu Ile Ile Gly Ser Phe Phe Met Leu
340 345 350

Asn Leu Val Leu Gly Val Leu Ser Gly Glu Phe Ala Lys Glu Arg Glu
355 360 365

Arg Val Glu Asn Arg Arg Ala Phe Leu Lys Leu Arg Arg Gln Gln Gln
370 375 380

Ile Glu Arg Glu Leu Asn Gly Tyr Met Glu Trp Ile Ser Lys Ala Glu
385 390 395 400

Glu Val Ile Leu Ala Glu Asp Glu Thr Asp Gly Glu Gln Arg His Pro
405 410 415

Phe Asp Gly Ala Leu Arg Arg Thr Thr Ile Lys Lys Ser Lys Thr Asp

-continued

420	425	430
Leu Leu Asn Pro Glu Glu Ala Glu Asp Gln Leu Ala Asp Ile Ala Ser		
435	440	445
Val Gly Ser Pro Phe Ala Arg Ala Ser Ile Lys Ser Ala Lys Leu Glu		
450	455	460
Asn Ser Thr Phe Phe His Lys Lys Glu Arg Arg Met Arg Phe Tyr Ile		
465	470	475
Arg Arg Met Val Lys Thr Gln Ala Phe Tyr Trp Thr Val Leu Ser Leu		
485	490	495
Val Ala Leu Asn Thr Leu Cys Val Ala Ile Val His Tyr Asn Gln Pro		
500	505	510
Glu Trp Leu Ser Asp Phe Leu Tyr Tyr Ala Glu Phe Ile Phe Leu Gly		
515	520	525
Leu Phe Met Ser Glu Met Phe Ile Lys Met Tyr Gly Leu Gly Thr Arg		
530	535	540
Pro Tyr Phe His Ser Ser Phe Asn Cys Phe Asp Cys Gly Val Ile Ile		
545	550	555
Gly Ser Ile Phe Glu Val Ile Trp Ala Val Ile Lys Pro Gly Thr Ser		
565	570	575
Phe Gly Ile Ser Val Leu Arg Ala Leu Arg Leu Leu Arg Ile Phe Lys		
580	585	590
Val Thr Lys Tyr Trp Ala Ser Leu Arg Asn Leu Val Val Ser Leu Leu		
595	600	605
Asn Ser Met Lys Ser Ile Ile Ser Leu Leu Phe Leu Leu Phe Leu Phe		
610	615	620
Ile Val Val Phe Ala Leu Leu Gly Met Gln Leu Phe Gly Gly Gln Phe		
625	630	635
Asn Phe Asp Glu Gly Thr Pro Pro Thr Asn Phe Asp Thr Phe Pro Ala		
645	650	655
Ala Ile Met Thr Val Phe Gln Ile Leu Thr Gly Glu Asp Trp Asn Glu		
660	665	670
Val Met Tyr Asp Gly Ile Lys Ser Gln Gly Gly Val Gln Gly Gly Met		
675	680	685
Val Phe Ser Ile Tyr Phe Ile Val Leu Thr Leu Phe Gly Asn Tyr Thr		
690	695	700
Leu Leu Asn Val Phe Leu Ala Ile Ala Val Asp Asn Leu Ala Asn Ala		
705	710	715
Gln Glu Leu Thr Lys Asp Glu Gln Glu Glu Glu Ala Ala Asn Gln		
725	730	735
Lys Leu Ala Leu Gln Lys Ala Lys Glu Val Ala Glu Val Ser Pro Leu		
740	745	750
Ser Ala Ala Asn Met Ser Ile Ala Val Lys Glu Gln Gln Lys Asn Gln		
755	760	765
Lys Pro Ala Lys Ser Val Trp Glu Gln Arg Thr Ser Glu Met Arg Lys		
770	775	780
Gln Asn Leu Leu Ala Ser Arg Glu Ala Leu Tyr Asn Glu Met Asp Pro		
785	790	795
Asp Glu Arg Trp Lys Ala Ala Tyr Thr Arg His Leu Arg Pro Asp Met		
805	810	815
Lys Thr His Leu Asp Arg Pro Leu Val Val Asp Pro Gln Glu Asn Arg		
820	825	830

-continued

Asn Asn Asn Thr Asn Lys Ser Arg Ala Ala Glu Pro Thr Val Asp Gln
 835 840 845
 Arg Leu Gly Gln Gln Arg Ala Glu Asp Phe Leu Arg Lys Gln Ala Arg
 850 855 860
 Tyr His Asp Arg Ala Arg Asp Pro Ser Gly Ser Ala Gly Leu Asp Ala
 865 870 875 880
 Arg Arg Pro Trp Ala Gly Ser Gln Glu Ala Glu Leu Ser Arg Glu Gly
 885 890 895
 Pro Tyr Gly Arg Glu Ser Asp His His Ala Arg Glu Gly Ser Leu Glu
 900 905 910
 Gln Pro Gly Phe Trp Glu Gly Glu Ala Glu Arg Gly Lys Ala Gly Asp
 915 920 925
 Pro His Arg Arg His Val His Arg Gln Gly Gly Ser Arg Glu Ser Arg
 930 935 940
 Ser Gly Ser Pro Arg Thr Gly Ala Asp Gly Glu His Arg Arg His Arg
 945 950 955 960
 Ala His Arg Arg Pro Gly Glu Gly Pro Glu Asp Lys Ala Glu Arg
 965 970 975
 Arg Ala Arg His Arg Glu Gly Ser Arg Pro Ala Arg Gly Gly Glu Gly
 980 985 990
 Glu Gly Glu Gly Pro Asp Gly Glu Arg Arg Arg Arg His Arg His
 995 1000 1005
 Gly Ala Pro Ala Thr Tyr Glu Gly Asp Ala Arg Arg Glu Asp Lys Glu
 1010 1015 1020
 Arg Arg His Arg Arg Arg Lys Glu Asn Gln Gly Ser Gly Val Pro Val
 1025 1030 1035 1040
 Ser Gly Pro Asn Leu Ser Thr Thr Arg Pro Ile Gln Gln Asp Leu Gly
 1045 1050 1055
 Arg Gln Asp Pro Pro Leu Ala Glu Asp Ile Asp Asn Met Lys Asn Asn
 1060 1065 1070
 Lys Leu Ala Thr Ala Glu Ser Ala Ala Pro His Gly Ser Leu Gly His
 1075 1080 1085
 Ala Gly Leu Pro Gln Ser Pro Ala Lys Met Gly Asn Ser Thr Asp Pro
 1090 1095 1100
 Gly Pro Met Leu Ala Ile Pro Ala Met Ala Thr Asn Pro Gln Asn Ala
 1105 1110 1115 1120
 Ala Ser Arg Arg Thr Pro Asn Asn Pro Gly Asn Pro Ser Asn Pro Gly
 1125 1130 1135
 Pro Pro Lys Thr Pro Glu Asn Ser Leu Ile Val Thr Asn Pro Ser Gly
 1140 1145 1150
 Thr Gln Thr Asn Ser Ala Lys Thr Ala Arg Lys Pro Asp His Thr Thr
 1155 1160 1165
 Val Asp Ile Pro Pro Ala Cys Pro Pro Pro Leu Asn His Thr Val Val
 1170 1175 1180
 Gln Val Asn Lys Asn Ala Asn Pro Asp Pro Leu Pro Lys Lys Glu Glu
 1185 1190 1195 1200
 Glu Lys Lys Glu Glu Glu Asp Asp Arg Gly Glu Asp Gly Pro Lys
 1205 1210 1215
 Pro Met Pro Pro Tyr Ser Ser Met Phe Ile Leu Ser Thr Thr Asn Pro
 1220 1225 1230

-continued

Leu Arg Arg Leu Cys His Tyr Ile Leu Asn Leu Arg Tyr Phe Glu Met
 1235 1240 1245
 Cys Ile Leu Met Val Ile Ala Met Ser Ser Ile Ala Leu Ala Ala Glu
 1250 1255 1260
 Asp Pro Val Gln Pro Asn Ala Pro Arg Asn Asn Val Leu Arg Tyr Phe
 1265 1270 1275 1280
 Asp Tyr Val Phe Thr Gly Val Phe Thr Phe Glu Met Val Ile Lys Met
 1285 1290 1295
 Ile Asp Leu Gly Leu Val Leu His Gln Gly Ala Tyr Phe Arg Asp Leu
 1300 1305 1310
 Trp Asn Ile Leu Asp Phe Ile Val Val Ser Gly Ala Leu Val Ala Phe
 1315 1320 1325
 Ala Phe Thr Gly Asn Ser Lys Gly Lys Asp Ile Asn Thr Ile Lys Ser
 1330 1335 1340
 Leu Arg Val Leu Arg Val Leu Arg Pro Leu Lys Thr Ile Lys Arg Leu
 1345 1350 1355 1360
 Pro Lys Leu Lys Ala Val Phe Asp Cys Val Val Asn Ser Leu Lys Asn
 1365 1370 1375
 Val Phe Asn Ile Leu Ile Val Tyr Met Leu Phe Met Phe Ile Phe Ala
 1380 1385 1390
 Val Val Ala Val Gln Leu Phe Lys Gly Lys Phe Phe His Cys Thr Asp
 1395 1400 1405
 Glu Ser Lys Glu Phe Glu Lys Asp Cys Arg Gly Lys Tyr Leu Leu Tyr
 1410 1415 1420
 Glu Lys Asn Glu Val Lys Ala Arg Asp Arg Glu Trp Lys Lys Tyr Glu
 1425 1430 1435 1440
 Phe His Tyr Asp Asn Val Leu Trp Ala Leu Leu Thr Leu Phe Thr Val
 1445 1450 1455
 Ser Thr Gly Glu Gly Trp Pro Gln Val Leu Lys His Ser Val Asp Ala
 1460 1465 1470
 Thr Phe Glu Asn Gln Gly Pro Ser Pro Gly Tyr Arg Met Glu Met Ser
 1475 1480 1485
 Ile Phe Tyr Val Val Tyr Phe Val Val Phe Pro Phe Phe Val Asn
 1490 1495 1500
 Ile Phe Val Ala Leu Ile Ile Ile Thr Phe Gln Glu Gln Gly Asp Lys
 1505 1510 1515 1520
 Met Met Glu Glu Tyr Ser Leu Glu Lys Asn Glu Arg Ala Cys Ile Asp
 1525 1530 1535
 Phe Ala Ile Ser Ala Lys Pro Leu Thr Arg His Met Pro Gln Asn Lys
 1540 1545 1550
 Gln Ser Phe Gln Tyr Arg Met Trp Gln Phe Val Val Ser Pro Pro Phe
 1555 1560 1565
 Glu Tyr Thr Ile Met Ala Met Ile Ala Leu Asn Thr Ile Val Leu Met
 1570 1575 1580
 Met Lys Phe Tyr Gly Ala Ser Val Ala Tyr Glu Asn Ala Leu Arg Val
 1585 1590 1595 1600
 Phe Asn Ile Val Phe Thr Ser Leu Phe Ser Leu Glu Cys Val Leu Lys
 1605 1610 1615
 Val Met Ala Phe Gly Ile Leu Asn Tyr Phe Arg Asp Ala Trp Asn Ile
 1620 1625 1630
 Phe Asp Phe Val Thr Val Leu Gly Ser Ile Thr Asp Ile Leu Val Thr

-continued

1635	1640	1645
Glu Phe Gly Asn Asn Phe Ile Asn Leu Ser Phe Leu Arg Leu Phe Arg		
1650	1655	1660
Ala Ala Arg Leu Ile Lys Leu Leu Arg Gln Gly Tyr Thr Ile Arg Ile		
1665	1670	1675
Leu Leu Trp Thr Phe Val Gln Ser Phe Lys Ala Leu Pro Tyr Val Cys		
1685	1690	1695
Leu Leu Ile Ala Met Leu Phe Phe Ile Tyr Ala Ile Ile Gly Met Gln		
1700	1705	1710
Val Phe Gly Asn Ile Gly Ile Asp Val Glu Asp Glu Asp Ser Asp Glu		
1715	1720	1725
Asp Glu Phe Gln Ile Thr Glu His Asn Asn Phe Arg Thr Phe Phe Gln		
1730	1735	1740
Ala Leu Met Leu Leu Phe Arg Ser Ala Thr Gly Glu Ala Trp His Asn		
1745	1750	1755
Ile Met Leu Ser Cys Leu Ser Gly Lys Pro Cys Asp Lys Asn Ser Gly		
1765	1770	1775
Ile Leu Thr Arg Glu Cys Gly Asn Glu Phe Ala Tyr Phe Tyr Phe Val		
1780	1785	1790
Ser Phe Ile Phe Leu Cys Ser Phe Leu Met Leu Asn Leu Phe Val Ala		
1795	1800	1805
Val Ile Met Asp Asn Phe Glu Tyr Leu Thr Arg Asp Ser Ser Ile Leu		
1810	1815	1820
Gly Pro His His Leu Asp Glu Tyr Val Arg Val Trp Ala Glu Tyr Asp		
1825	1830	1835
Pro Ala Ala Cys Gly Arg Ile His Tyr Lys Asp Met Tyr Ser Leu Leu		
1845	1850	1855
Arg Val Ile Ser Pro Pro Leu Gly Leu Gly Lys Lys Cys Pro His Arg		
1860	1865	1870
Val Ala Cys Lys Arg Leu Leu Arg Met Asp Leu Pro Val Ala Asp Asp		
1875	1880	1885
Asn Thr Val His Phe Asn Ser Thr Leu Met Ala Leu Ile Arg Thr Ala		
1890	1895	1900
Leu Asp Ile Lys Ile Ala Lys Gly Gly Ala Asp Lys Gln Gln Met Asp		
1905	1910	1915
Ala Glu Leu Arg Lys Glu Met Met Ala Ile Trp Pro Asn Leu Ser Gln		
1925	1930	1935
Lys Thr Leu Asp Leu Leu Val Thr Pro His Lys Ser Thr Asp Leu Thr		
1940	1945	1950
Val Gly Lys Ile Tyr Ala Ala Met Met Ile Met Glu Tyr Tyr Arg Gln		
1955	1960	1965
Ser Lys Ala Lys Lys Leu Gln Ala Met Arg Glu Glu Gln Asp Arg Thr		
1970	1975	1980
Pro Leu Met Phe Gln Arg Met Glu Pro Pro Ser Pro Thr Gln Glu Gly		
1985	1990	1995
Gly Pro Gly Gln Asn Ala Leu Pro Ser Thr Gln Leu Asp Pro Gly Gly		
2005	2010	2015
Ala Leu Met Ala His Glu Ser Gly Leu Lys Glu Ser Pro Ser Trp Val		
2020	2025	2030
Thr Gln Arg Ala Gln Glu Met Phe Gln Lys Thr Gly Thr Trp Ser Pro		
2035	2040	2045

-continued

Glu Gln Gly Pro Pro Thr Asp Met Pro Asn Ser Gln Pro Asn Ser Gln
2050 2055 2060

Ser Val Glu Met Arg Glu Met Gly Arg Asp Gly Tyr Ser Asp Ser Glu
2065 2070 2075 2080

His Tyr Leu Pro Met Glu Gly Gln Gly Arg Ala Ala Ser Met Pro Arg
2085 2090 2095

Leu Pro Ala Glu Asn Gln Arg Arg Gly Arg Pro Arg Gly Asn Asn
2100 2105 2110

Leu Ser Thr Ile Ser Asp Thr Ser Pro Met Lys Arg Ser Ala Ser Val
2115 2120 2125

Leu Gly Pro Lys Ala Arg Arg Leu Asp Asp Tyr Ser Leu Glu Arg Val
2130 2135 2140

Pro Pro Glu Glu Asn Gln Arg His His Gln Arg Arg Arg Asp Arg Ser
2145 2150 2155 2160

His Arg Ala Ser Glu Arg Ser Leu Gly Arg Tyr Thr Asp Val Asp Thr
2165 2170 2175

Gly Leu Gly Thr Asp Leu Ser Met Thr Thr Gln Ser Gly Asp Leu Pro
2180 2185 2190

Ser Lys Glu Arg Asp Gln Glu Arg Gly Arg Pro Lys Asp Arg Lys His
2195 2200 2205

Arg Gln His Pro Pro Pro Pro
2210 2215 2220

Asp Lys Asp Arg Tyr Ala Gln Glu Arg Pro Asp His Gly Arg Ala Arg
2225 2230 2235 2240

Ala Arg Asp Gln Arg Trp Ser Arg Ser Pro Ser Glu Gly Arg Glu His
2245 2250 2255

Met Ala His Arg Gln
2260

<210> SEQ ID NO 7
<211> LENGTH: 7926
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 7

ggggtcaccc	cctctgcggg	catgtcccgaa	gcccgcgtcc	ccggctcgcc	cggggcgtcc	60
cgctcgttag	ccccccggc	tgaggccggg	cgccgcggcc	cgatgggctg	ggccgtggag	120
cgtctccgca	gtgcaggtg	cagccgcgt	ccccacagcc	cgccgcgcct	gagcagcggc	180
ggcggtggtg	gtgggtggtg	tggcggtggc	cgtggcttcc	gcagcgcctcg	ccgcacgaac	240
agcagccgccc	gcgttaaccgg	aggccctttg	ctctttgcag	aatggccgc	tttggagacg	300
agatgccggg	ccgctacggc	gcaggcggag	gaggctcagg	gccggccgcc	gggggtggtc	360
tgggcgcgc	gggcggccga	ggagccgggg	gcagcggca	gggcgggcag	cccgaggcgc	420
agaggatgt	caagcagtgc	atggcgca	gagcgcggac	catggccctc	tacaacccca	480
tccctgtccg	ccagaactgc	ctcacggta	accgctccct	gttcccttc	agtgaagaca	540
acgtggtgag	aaaatacgc	aaaaagatca	ctgaatggcc	tcccttc	gatcatgtatcc	600
tggccaccat	cattgcta	ac	tgcatacg	tggccctgga	gcagcacctc	660
acaagacacc	catgtctgag	cggctggatg	acacggaa	cc	atacttcatt	720
gttttggc	tggaattaag	atcgtggctc	ttggctttgc	cttccacaaa	ggctcc	780

-continued

tggagaatgg ctggaacgtc atggactttg tcgtggtgct aacaggcata ttggccactg	840
tcgggacgga gtttgccta cggacactga gggcggttcg tgtgctgcgg ccactcaagc	900
tgggtctgg aatcccaagt ttacaagtgc tcctgaagtc aatcatgaag gcgtatgtcc	960
ctctgctgca gatcgccctc ctcctgtttt ttgcaatcct tattttgca atcataggg	1020
tagaatttta tatggggaaa tttcataccca cctgcttga agaggggaca gacgacatcc	1080
agggtgagtc gccagctcg tggtggacag aggagcctgc ccgcacccgc cccaaacggga	1140
ccaaatgtca gccgtactgg gaagggccca acaacggcat cactcagttc gacaacatcc	1200
tgtttgctgt gctcaactgtt ttccagtgc taaccatggta aggctggact gatctctct	1260
acaatagcaa cgatgcctca gggAACACTT ggaactgggtt gtacttcatc cccctcatca	1320
tcatcgccctc ctttttatg ctgaacccgtt ttctgggtgt gctgtctggg gagtttgc	1380
aagaaaaggga acgtgtggag aaccggagag cttttctgaa gtcagaaga caacagcaga	1440
ttgaacgtga gctcaatggta tacatggagt ggatctcaaa agcagaagag gtgatttttgc	1500
cggaggacga gacagacgtg gagcagaggg acccttttga tggagctt cggagagcta	1560
ctctgaagaa aagcaagaca gacctgtca accctgagga ggccgaggac cagctggctg	1620
acatcgccgtc tgggggtct cccttcgcaca gagccagcat caaaagtgc aagctggaga	1680
actcgacttt ttttcacaag aaggagagaa gaatgcgtt ctacatccgc cgcattgtca	1740
aaactcaggc cttctactgg accgtgtca gtctggtagc cctcaaacacg ctgtgogtgc	1800
ccattgtcca ctacaaccag cccgagtgcc totccgactt cctctactat gcagaattca	1860
ttttcttagg actctttatg tccgaaatgt ttataaaaat gtacgggctc gggacgggc	1920
cttacttcca cttttccttc aactgccttg actgtgggg catcattggg agcatcttg	1980
aagtcatctg ggctgtcatc aaacctggta catccttgg aatcagtgtt ttacgagctc	2040
tcaggttatt gcgaattttc aaagtcaaca agtactgggc atctctcaga aacctgggt	2100
tctccctctt caactccatg aagtccatca taagtctgtt gttccctctt ttccctgttca	2160
tcgttgtctt tgeccctctg gggatgcgc tggtgggtt ccaattcaat ttgtatgggg	2220
ggactccgcc caccacccgc gacacttttgc cagcageaat aatgactgtt tttcagatcc	2280
tgactggcga ggattggaa gaggtcatgt atgatggat caagtctcag gggggcgtgc	2340
aggggccgcattt ggtgttctcc atctacttca tcgtccctcac cctcttcggg aactacaccc	2400
tgctgaacgt gttttggcc atcgcaactgg acaacccgtc caatgecccg gagctccacca	2460
aggatgagca agaagaggaa gaggcagccaa atcagaagct ggctctacag aaagccaaagg	2520
aggtagcaga agtgagtcgg ctgtctgcgc ccaacatgtc catageccgtg aaggagcagc	2580
agaagaacca gaagcctacc aagtccgtgt gggagcagcg caccagcagc atgcgcacgc	2640
agaacactgtt ggcacggccg gaggcgctgt acggggacgc ggctgagcgc tggcccaccc	2700
cctacgcgcgg cccgctgcgg cccgacgtga agactcactt ggaccggccg cttgtgggt	2760
acccgcagga gaaaaatggaa aacaacacca acaagagccg tgccggcagaa ggcgtgcgc	2820
cgaccgcgcg gccccgcgaa agcgccgcgc accccgcacgc gggccgcgc tggccggca	2880
gcctcgccgc cgccccctggaa cgagaggccgc cgtatgggcg cgagagcagc cctcagcagc	2940
gcgcgcacgc gccccgcgcgaa gaggcagccac cctgggacgc ggacaccgag cgcccaaaag	3000
ccggggacgc gccccgcgcgac cacacgcacc gacccgtggc ggagggcagc cctcgccgc	3060

-continued

accacgccc	ccggccggccc	ggggacgaac	ccgacgacag	gccagagcgc	aggccacgtc	3120
cccgcgacgc	cactaggccg	gcccgcgcgt	cagacggcga	aggcgacgt	ggggagcga	3180
agcggcgaca	ccgacacggg	ccggccggccc	acgacgacag	ggagcgcaga	caccgacgca	3240
gaaaagagaa	ccagggctct	ggggtccccg	tgtctggtcc	caacttgtcc	accaccaggc	3300
caatccagca	ggacctgggc	cgccaggacc	tgccgctggc	tgaggatctt	gacaacatga	3360
agaataacaa	gctggccacc	gggaaacctg	ccagtcctca	cgacagcctg	ggccacagcg	3420
gccttccccc	cagccccgcc	aaaatcgga	acagcaccaa	ccctggtccc	gccttggcca	3480
ccaaatcccc	gaacgctgcc	agccgcagg	cgcccaacaa	tccggcaac	ccgtccaaacc	3540
ccggccccc	caagactccc	gagaacagcc	ttatcgtcac	caacccca	agcaccaggc	3600
ccaaactcage	gaagactgcc	aggaaacccg	agcacatggc	ggtggagatc	cccccagcct	3660
gcccggccct	caaccacacc	gtggtccaag	taaacaaaaa	cgccaaaccc	gaccactgc	3720
caaagaaa	ggaagagaag	aaagaggaag	aggaggcaga	ccccggggag	gacggcccaa	3780
agcccatgcc	gcctacagc	tccatgttca	tcctctccac	caccaacccc	cttgcggc	3840
tgtgccatta	cattctgaac	ctgcgcact	tgcagatgtg	cattctcatg	gtcattggca	3900
tgagcagcat	cgetctggcc	gcggaggacc	cggtgcagcc	caacgcaccc	cgcaacaatg	3960
tgctgcgata	ttttgactat	gttttacag	gctgtttac	ctttagatg	gtgatcaaga	4020
tgatcgacct	gggcctcg	ttgcattcagg	gggcctattt	ccgtgaccta	tggAACATT	4080
tggacttcat	agtggtcagt	ggggccctgg	tggcccttgc	cttcactggc	aatagoaaag	4140
gaaaggacat	taacaccatt	aagtctctcc	gagtcctccg	ggtgctacga	cctctaaaga	4200
ccatcaageg	gctgccaag	ctgaaggccg	tgtttgactg	cgtggtaac	tgcgtcaaga	4260
acgtttcaa	catoctcatt	gtgtacatgc	tgttcatgtt	catcttgc	gtggtagccg	4320
tgcagctctt	caagggcaag	tttttccact	gcacggacga	gtccaaggag	tgcagagag	4380
actgtcgagg	caagtaccc	ctgtacgaga	agaacgaagt	gaaggccccg	gaccgogagt	4440
ggaagaagta	cgttccac	tacgacaac	tgctctggc	tctgctaca	ctttcacag	4500
tgtccacagg	agagggctgg	ccacagg	tcaagcactc	agtggatgt	actttgaga	4560
accaggccc	cagccccggg	tacccatgg	agatgtccat	cttctacgtg	gtgtacttt	4620
tagtgtttcc	cttcttctt	gtcaatatct	tgcgtgcctt	gatcatcatc	acctttcagg	4680
agcaggggaga	caagatgtat	gaagagtaca	gtttggagaa	aatgagagg	gcctgcac	4740
actttgccat	cagtgcac	ccgctgacca	ggcacatg	ccagaacaag	cagagttcc	4800
agtatcgaat	gtggcagttc	gtcggtgtcc	cacccttga	gtacaccatc	atggccat	4860
tgcgtctcaa	caccatcg	ctaatgtga	agttctacgg	agcctcggtg	gcctatgaaa	4920
acgeccctcg	cgtgttcaac	attgtctca	cctccctt	ctctctcgaa	tgtgtgc	4980
aagtcatggc	ttttgggatt	ctgaattatt	tccgcgtatc	ctggaaacatc	tgcacttt	5040
tgactgttct	gggcagc	acagacatcc	tgcgtacccg	gtttgggaa	aacttcatca	5100
acctgagctt	tctccgcctc	ttccgtgtc	ccagactcat	caaactcctc	cgccagggtt	5160
acaccatccg	cattctcc	tggacttcc	tgcagtc	caaggccct	ccttatgtct	5220
gtctgctgat	cgccatg	ttcttcatct	acgccc	cggtatgc	gtgtttggca	5280
acatcgccat	tgacgggaa	gacgaggaca	gcgacgagga	tgaattccaa	atcacggac	5340

-continued

acaataactt	ccggacacctc	ttccaagcgc	tcatgcttct	cttccggaga	gccacaggag	5400
aagcgtggca	caacatcatg	ctctcctgcc	tcagcggaa	gccgtgcgt	aagaactctg	5460
ggatcctcac	ggcagactgc	ggcaacgagt	tcgcgtactt	ctactttgtc	tcgttcatct	5520
tcctgtgctc	gtttctgtat	ctgaatctct	ttgtcgctgt	catcatggac	aacttcgagt	5580
acctcaccgg	agattcttcc	atcctggcc	ccatcactt	ggatgaatac	gtgcgtgtct	5640
gggcagagta	tgaccctgt	gcctgcggcc	gcattcacta	taaggacatg	tacagtttat	5700
tgcgagtaat	atcgccccct	ctcggcttag	gcaagaaatg	tcctcatagg	gttgcttgca	5760
agagactctt	gcggatggac	ctgcccgttag	catatgacaa	cacagttcac	ttcaactcca	5820
ccctgatggc	tctgatccga	accgccttg	atatcaaata	tgccaagggt	ggagctgaca	5880
agcagcagat	ggatgcagag	cttcgcaagg	agatgtatggc	catttggccc	aacctgtctc	5940
agaagaccct	ggatctgtcg	gtcacaccc	acaagtccac	ggacctgaca	gtgggtaaga	6000
tctacgcage	catatgtatc	atggagtatt	accggcagag	caaggccaa	aaactgcagg	6060
ccatgcgaga	ggagcagaac	cgAACACAC	tcatgttcca	gcccgtggag	cctccatcac	6120
caacacagga	gggaggagcc	agccagaatg	cccttccttc	cactcagctg	gaccaggag	6180
gaggcctgat	gggtcacgaa	ggggcatga	aggagagccc	gtcctgggtg	acccagcggg	6240
cccaagagat	gttccagaag	actggcacct	ggagcccaga	gcccggccaa	cccatogaca	6300
tgcctaacag	ccagcccaac	tcccagtctg	tggagatgcg	agaaatgggt	actgtatggct	6360
actcagacag	cgaacactac	ctcccccattg	aaggccagac	caggggccccc	tccatgcccc	6420
gcctccca	agagaaccag	aggagaaggg	gcccggccacg	tggaaatgac	ctcagtacca	6480
tctctgatac	cagccccatg	aagcgctca	cctctgtgt	gggacccaaa	gcccggcgac	6540
tggatgacta	ctcaactagag	cgggtacccg	ctgaggagaa	ccaaacggta	caccaacgtc	6600
gcccggaccc	tggccacccgc	acctctgagc	gtctctggg	ccgatacact	gatgtggaca	6660
caggtttggg	gacagatctg	agcatgacca	ccaaatcggg	tgacctgccc	tccaaagatc	6720
gggaccagga	ccggggccgca	ccaaaggacc	ggaagcatcg	gccacaccac	caccaccacc	6780
accatcatca	tcatcccccg	gccccccgacc	gggatecgcta	cgcacaggag	cgccgggaca	6840
ccggccgggc	gccccggccgg	gagcagecgct	gttcccgctc	gcccagcgaa	gttcgggagc	6900
acacgacaca	cagacagggc	agtatcccg	tcaatggaa	cccaatcccc	tctacgtccg	6960
gcaccacgac	gcggcggegc	ggccgcggcc	agctccccca	gacccctgc	accccgccgc	7020
cgcgggtgtc	ctactcgccc	gtcccgccca	ggcctgcggc	gcccaggatg	gcgggtcccc	7080
cageggccccc	tggcggtctg	ccgggggggt	gcccggcgct	gcccggcg	cctcgccatg	7140
cgcctgaggg	cccgccggccg	cgccgtgcgc	actacaccga	gcccggacagc	ccgcggcgac	7200
cgcgggggg	cgegcacgat	cccgccggcc	ggtcgcccag	gactccgcgc	gcccggggct	7260
cgcggttacc	cgccacggcc	cgccggcgct	ccaatggta	ctacgggggg	cacggccgc	7320
cgcggccggc	cacagccccgc	agggggccgc	acgacggcgt	cagcgagacg	gaggacgact	7380
ggtgctaagc	cgcggccccc	cccgccggcc	ccgtaaatgt	cgccgcact	gcaaatgata	7440
aacaataacg	ttttaaaaaa	agaaaacccg	ggagaaaaac	aaaaatagct	tctattgtat	7500
agttttatcg	tctcgattga	atcttgcgt	ttcccggtga	atgaaacgca	ggtggctgca	7560
agccggaagg	acggctgcac	gccccacaca	cttctcagac	ccacgggagt	gttgcgggc	7620

-continued

cacgccaaac ttattactgc ctgcagagat caactacaag aaaacaacac taacagttt	7680
taaaaggaca aaaaaaaaaatt gagaaaaagt tttttctga tatttggtcc tgcttggaaat	7740
aacaaaagaa aaaagaaaaa attaccacca tcactgattc ccttgctct tttttcctt	7800
tatctttta aaaatttttt ttcctttttt tccttatcttg tttgaaaaac gtgggcttgg	7860
gactgtgaaa tattgcatga cattaaaaag aacaaaaaaaaa taataaaaaac gaaaaaaaaactt	7920
gaatcc	7926

<210> SEQ_ID NO 8
<211> LENGTH: 2368
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 8

Met Ala Arg Phe Gly Asp Glu Met Pro Gly Arg Tyr Gly Ala Gly Gly	
1 5 10 15	
Gly Gly Ser Gly Pro Ala Ala Gly Val Val Val Gly Ala Ala Gly Gly	
20 25 30	
Arg Gly Ala Gly Gly Ser Arg Gln Gly Gly Gln Pro Gly Ala Gln Arg	
35 40 45	
Met Tyr Lys Gln Ser Met Ala Gln Arg Ala Arg Thr Met Ala Leu Tyr	
50 55 60	
Asn Pro Ile Pro Val Arg Gln Asn Cys Leu Thr Val Asn Arg Ser Leu	
65 70 75 80	
Phe Leu Phe Ser Glu Asp Asn Val Val Arg Lys Tyr Ala Lys Lys Ile	
85 90 95	
Thr Glu Trp Pro Pro Phe Glu Tyr Met Ile Leu Ala Thr Ile Ile Ala	
100 105 110	
Asn Cys Ile Val Leu Ala Leu Glu Gln His Leu Pro Asp Asp Asp Lys	
115 120 125	
Thr Pro Met Ser Glu Arg Leu Asp Asp Thr Glu Pro Tyr Phe Ile Gly	
130 135 140	
Ile Phe Cys Phe Glu Ala Gly Ile Lys Ile Val Ala Leu Gly Phe Ala	
145 150 155 160	
Phe His Lys Gly Ser Tyr Leu Arg Asn Gly Trp Asn Val Met Asp Phe	
165 170 175	
Val Val Val Leu Thr Gly Ile Leu Ala Thr Val Gly Thr Glu Phe Asp	
180 185 190	
Leu Arg Thr Leu Arg Ala Val Arg Val Leu Arg Pro Leu Lys Leu Val	
195 200 205	
Ser Gly Ile Pro Ser Leu Gln Val Val Leu Lys Ser Ile Met Lys Ala	
210 215 220	
Met Ile Pro Leu Leu Gln Ile Gly Leu Leu Leu Phe Phe Ala Ile Leu	
225 230 235 240	
Ile Phe Ala Ile Ile Gly Leu Glu Phe Tyr Met Gly Lys Phe His Thr	
245 250 255	
Thr Cys Phe Glu Glu Gly Thr Asp Asp Ile Gln Gly Glu Ser Pro Ala	
260 265 270	
Pro Cys Gly Thr Glu Glu Pro Ala Arg Thr Cys Pro Asn Gly Thr Lys	
275 280 285	
Cys Gln Pro Tyr Trp Glu Gly Pro Asn Asn Gly Ile Thr Gln Phe Asp	

-continued

290	295	300
Asn Ile Leu Phe Ala Val	Leu Thr Val Phe Gln Cys	Ile Thr Met Glu
305	310	315
320		
Gly Trp Thr Asp Leu Leu Tyr Asn Ser Asn Asp Ala Ser Gly Asn Thr		
325	330	335
Trp Asn Trp Leu Tyr Phe Ile Pro Leu Ile Ile Gly Ser Phe Phe		
340	345	350
Met Leu Asn Leu Val Leu Gly Val Leu Ser Gly Glu Phe Ala Lys Glu		
355	360	365
Arg Glu Arg Val Glu Asn Arg Arg Ala Phe Leu Lys Leu Arg Arg Gln		
370	375	380
Gln Gln Ile Glu Arg Glu Leu Asn Gly Tyr Met Glu Trp Ile Ser Lys		
385	390	395
400		
Ala Glu Glu Val Ile Leu Ala Glu Asp Glu Thr Asp Val Glu Gln Arg		
405	410	415
His Pro Phe Asp Gly Ala Leu Arg Arg Ala Thr Leu Lys Lys Ser Lys		
420	425	430
Thr Asp Leu Leu Asn Pro Glu Glu Ala Glu Asp Gln Leu Ala Asp Ile		
435	440	445
Ala Ser Val Gly Ser Pro Phe Ala Arg Ala Ser Ile Lys Ser Ala Lys		
450	455	460
Leu Glu Asn Ser Thr Phe Phe His Lys Lys Glu Arg Arg Met Arg Phe		
465	470	475
480		
Tyr Ile Arg Arg Met Val Lys Thr Gln Ala Phe Tyr Trp Thr Val Leu		
485	490	495
Ser Leu Val Ala Leu Asn Thr Leu Cys Val Ala Ile Val His Tyr Asn		
500	505	510
Gln Pro Glu Trp Leu Ser Asp Phe Leu Tyr Tyr Ala Glu Phe Ile Phe		
515	520	525
Leu Gly Leu Phe Met Ser Glu Met Phe Ile Lys Met Tyr Gly Leu Gly		
530	535	540
Thr Arg Pro Tyr Phe His Ser Ser Phe Asn Cys Phe Asp Cys Gly Val		
545	550	555
560		
Ile Ile Gly Ser Ile Phe Glu Val Ile Trp Ala Val Ile Lys Pro Gly		
565	570	575
Thr Ser Phe Gly Ile Ser Val Leu Arg Ala Leu Arg Leu Leu Arg Ile		
580	585	590
Phe Lys Val Thr Lys Tyr Trp Ala Ser Leu Arg Asn Leu Val Val Ser		
595	600	605
Leu Leu Asn Ser Met Lys Ser Ile Ile Ser Leu Leu Phe Leu Leu Phe		
610	615	620
Leu Phe Ile Val Val Phe Ala Leu Leu Gly Met Gln Leu Phe Gly Gly		
625	630	635
640		
Gln Phe Asn Phe Asp Glu Gly Thr Pro Pro Thr Asn Phe Asp Thr Phe		
645	650	655
Pro Ala Ala Ile Met Thr Val Phe Gln Ile Leu Thr Gly Glu Asp Trp		
660	665	670
Asn Glu Val Met Tyr Asp Gly Ile Lys Ser Gln Gly Gly Val Gln Gly		
675	680	685
Gly Met Val Phe Ser Ile Tyr Phe Ile Val Leu Thr Leu Phe Gly Asn		
690	695	700

-continued

Tyr Thr Leu Leu Asn Val Phe Leu Ala Ile Ala Val Asp Asn Leu Ala
705 710 715 720

Asn Ala Gln Glu Leu Thr Lys Asp Glu Gln Glu Glu Glu Ala Ala
725 730 735

Asn Gln Lys Leu Ala Leu Gln Lys Ala Lys Glu Val Ala Glu Val Ser
740 745 750

Pro Leu Ser Ala Ala Asn Met Ser Ile Ala Val Lys Glu Gln Gln Lys
755 760 765

Asn Gln Lys Pro Thr Lys Ser Val Trp Glu Gln Arg Thr Ser Glu Met
770 775 780

Arg Lys Gln Asn Leu Leu Ala Ser Arg Glu Ala Leu Tyr Gly Asp Ala
785 790 795 800

Ala Glu Arg Trp Pro Thr Pro Tyr Ala Arg Pro Leu Arg Pro Asp Val
805 810 815

Lys Thr His Leu Asp Arg Pro Leu Val Val Asp Pro Gln Glu Asn Arg
820 825 830

Asn Asn Asn Thr Asn Lys Ser Arg Ala Pro Glu Ala Leu Arg Pro Thr
835 840 845

Ala Arg Pro Arg Glu Ser Ala Arg Asp Pro Asp Ala Arg Arg Ala Trp
850 855 860

Pro Gly Ser Pro Glu Arg Ala Pro Gly Arg Glu Gly Pro Tyr Gly Arg
865 870 875 880

Glu Ser Glu Pro Gln Gln Arg Glu His Ala Pro Pro Arg Glu His Ala
885 890 895

Pro Trp Asp Ala Asp Thr Glu Arg Ala Lys Ala Gly Asp Ala Pro Arg
900 905 910

Arg His Thr His Arg Pro Val Ala Glu Gly Glu Pro Arg Arg His His
915 920 925

Ala Arg Arg Arg Pro Gly Asp Glu Pro Asp Asp Arg Pro Glu Arg Arg
930 935 940

Pro Arg Pro Arg Asp Ala Thr Arg Pro Ala Arg Ala Ala Asp Gly Glu
945 950 955 960

Gly Asp Asp Gly Glu Arg Lys Arg Arg His Arg His Gly Pro Pro Ala
965 970 975

His Asp Asp Arg Glu Arg Arg His Arg Arg Arg Lys Glu Asn Gln Gly
980 985 990

Ser Gly Val Pro Val Ser Gly Pro Asn Leu Ser Thr Thr Arg Pro Ile
995 1000 1005

Gln Gln Asp Leu Gly Arg Gln Asp Leu Pro Leu Ala Glu Asp Leu Asp
1010 1015 1020

Asn Met Lys Asn Asn Lys Leu Ala Thr Gly Glu Pro Ala Ser Pro His
1025 1030 1035 1040

Asp Ser Leu Gly His Ser Gly Leu Pro Pro Ser Pro Ala Lys Ile Gly
1045 1050 1055

Asn Ser Thr Asn Pro Gly Pro Ala Leu Ala Thr Asn Pro Gln Asn Ala
1060 1065 1070

Ala Ser Arg Arg Thr Pro Asn Asn Pro Gly Asn Pro Ser Asn Pro Gly
1075 1080 1085

Pro Pro Lys Thr Pro Glu Asn Ser Leu Ile Val Thr Asn Pro Ser Ser
1090 1095 1100

-continued

Thr Gln Pro Asn Ser Ala Lys Thr Ala Arg Lys Pro Glu His Met Ala
 1105 1110 1115 1120
 Val Glu Ile Pro Pro Ala Cys Pro Pro Leu Asn His Thr Val Val Gln
 1125 1130 1135
 Val Asn Lys Asn Ala Asn Pro Asp Pro Leu Pro Lys Lys Glu Glu Glu
 1140 1145 1150
 Lys Lys Glu Glu Glu Ala Asp Pro Gly Glu Asp Gly Pro Lys Pro
 1155 1160 1165
 Met Pro Pro Tyr Ser Ser Met Phe Ile Leu Ser Thr Thr Asn Pro Leu
 1170 1175 1180
 Arg Arg Leu Cys His Tyr Ile Leu Asn Leu Arg Tyr Phe Glu Met Cys
 1185 1190 1195 1200
 Ile Leu Met Val Ile Ala Met Ser Ser Ile Ala Leu Ala Ala Glu Asp
 1205 1210 1215
 Pro Val Gln Pro Asn Ala Pro Arg Asn Asn Val Leu Arg Tyr Phe Asp
 1220 1225 1230
 Tyr Val Phe Thr Gly Val Phe Thr Phe Glu Met Val Ile Lys Met Ile
 1235 1240 1245
 Asp Leu Gly Leu Val Leu His Gln Gly Ala Tyr Phe Arg Asp Leu Trp
 1250 1255 1260
 Asn Ile Leu Asp Phe Ile Val Val Ser Gly Ala Leu Val Ala Phe Ala
 1265 1270 1275 1280
 Phe Thr Gly Asn Ser Lys Gly Lys Asp Ile Asn Thr Ile Lys Ser Leu
 1285 1290 1295
 Arg Val Leu Arg Val Leu Arg Pro Leu Lys Thr Ile Lys Arg Leu Pro
 1300 1305 1310
 Lys Leu Lys Ala Val Phe Asp Cys Val Val Asn Ser Leu Lys Asn Val
 1315 1320 1325
 Phe Asn Ile Leu Ile Val Tyr Met Leu Phe Met Phe Ile Phe Ala Val
 1330 1335 1340
 Val Ala Val Gln Leu Phe Lys Gly Lys Phe Phe His Cys Thr Asp Glu
 1345 1350 1355 1360
 Ser Lys Glu Phe Glu Arg Asp Cys Arg Gly Lys Tyr Leu Leu Tyr Glu
 1365 1370 1375
 Lys Asn Glu Val Lys Ala Arg Asp Arg Glu Trp Lys Lys Tyr Glu Phe
 1380 1385 1390
 His Tyr Asp Asn Val Leu Trp Ala Leu Leu Thr Leu Phe Thr Val Ser
 1395 1400 1405
 Thr Gly Glu Gly Trp Pro Gln Val Leu Lys His Ser Val Asp Ala Thr
 1410 1415 1420
 Phe Glu Asn Gln Gly Pro Ser Pro Gly Tyr Arg Met Glu Met Ser Ile
 1425 1430 1435 1440
 Phe Tyr Val Val Tyr Phe Val Val Phe Pro Phe Phe Phe Val Asn Ile
 1445 1450 1455
 Phe Val Ala Leu Ile Ile Ile Thr Phe Gln Glu Gln Gly Asp Lys Met
 1460 1465 1470
 Met Glu Glu Tyr Ser Leu Glu Lys Asn Glu Arg Ala Cys Ile Asp Phe
 1475 1480 1485
 Ala Ile Ser Ala Lys Pro Leu Thr Arg His Met Pro Gln Asn Lys Gln
 1490 1495 1500
 Ser Phe Gln Tyr Arg Met Trp Gln Phe Val Val Ser Pro Pro Phe Glu

-continued

1505	1510	1515	1520
Tyr Thr Ile Met Ala Met Ile Ala Leu Asn Thr Ile Val Leu Met Met			
1525	1530	1535	
Lys Phe Tyr Gly Ala Ser Val Ala Tyr Glu Asn Ala Leu Arg Val Phe			
1540	1545	1550	
Asn Ile Val Phe Thr Ser Leu Phe Ser Leu Glu Cys Val Leu Lys Val			
1555	1560	1565	
Met Ala Phe Gly Ile Leu Asn Tyr Phe Arg Asp Ala Trp Asn Ile Phe			
1570	1575	1580	
Asp Phe Val Thr Val Leu Gly Ser Ile Thr Asp Ile Leu Val Thr Glu			
1585	1590	1595	1600
Phe Gly Asn Asn Phe Ile Asn Leu Ser Phe Leu Arg Leu Phe Arg Ala			
1605	1610	1615	
Ala Arg Leu Ile Lys Leu Leu Arg Gln Gly Tyr Thr Ile Arg Ile Leu			
1620	1625	1630	
Leu Trp Thr Phe Val Gln Ser Phe Lys Ala Leu Pro Tyr Val Cys Leu			
1635	1640	1645	
Leu Ile Ala Met Leu Phe Phe Ile Tyr Ala Ile Ile Gly Met Gln Val			
1650	1655	1660	
Phe Gly Asn Ile Gly Ile Asp Gly Glu Asp Ser Asp Glu Asp			
1665	1670	1675	1680
Glu Phe Gln Ile Thr Glu His Asn Asn Phe Arg Thr Phe Phe Gln Ala			
1685	1690	1695	
Leu Met Leu Leu Phe Arg Ser Ala Thr Gly Glu Ala Trp His Asn Ile			
1700	1705	1710	
Met Leu Ser Cys Leu Ser Gly Lys Pro Cys Asp Lys Asn Ser Gly Ile			
1715	1720	1725	
Leu Thr Ala Asp Cys Gly Asn Glu Phe Ala Tyr Phe Tyr Phe Val Ser			
1730	1735	1740	
Phe Ile Phe Leu Cys Ser Phe Leu Met Leu Asn Leu Phe Val Ala Val			
1745	1750	1755	1760
Ile Met Asp Asn Phe Glu Tyr Leu Thr Arg Asp Ser Ser Ile Leu Gly			
1765	1770	1775	
Pro His His Leu Asp Glu Tyr Val Arg Val Trp Ala Glu Tyr Asp Pro			
1780	1785	1790	
Ala Ala Cys Gly Arg Ile His Tyr Lys Asp Met Tyr Ser Leu Leu Arg			
1795	1800	1805	
Val Ile Ser Pro Pro Leu Gly Leu Gly Lys Lys Cys Pro His Arg Val			
1810	1815	1820	
Ala Cys Lys Arg Leu Leu Arg Met Asp Leu Pro Val Ala Asp Asp Asn			
1825	1830	1835	1840
Thr Val His Phe Asn Ser Thr Leu Met Ala Leu Ile Arg Thr Ala Leu			
1845	1850	1855	
Asp Ile Lys Ile Ala Lys Gly Ala Asp Lys Gln Gln Met Asp Ala			
1860	1865	1870	
Glu Leu Arg Lys Glu Met Met Ala Ile Trp Pro Asn Leu Ser Gln Lys			
1875	1880	1885	
Thr Leu Asp Leu Leu Val Thr Pro His Lys Ser Thr Asp Leu Thr Val			
1890	1895	1900	
Gly Lys Ile Tyr Ala Ala Met Met Ile Met Glu Tyr Tyr Arg Gln Ser			
1905	1910	1915	1920

-continued

Lys Ala Lys Lys Leu Gln Ala Met Arg Glu Glu Gln Asn Arg Thr Pro
 1925 1930 1935
 Leu Met Phe Gln Arg Met Glu Pro Pro Ser Pro Thr Gln Glu Gly Gly
 1940 1945 1950
 Pro Ser Gln Asn Ala Leu Pro Ser Thr Gln Leu Asp Pro Gly Gly
 1955 1960 1965
 Leu Met Ala His Glu Gly Gly Met Lys Glu Ser Pro Ser Trp Val Thr
 1970 1975 1980
 Gln Arg Ala Gln Glu Met Phe Gln Lys Thr Gly Thr Trp Ser Pro Glu
 1985 1990 1995 2000
 Arg Gly Pro Pro Ile Asp Met Pro Asn Ser Gln Pro Asn Ser Gln Ser
 2005 2010 2015
 Val Glu Met Arg Glu Met Gly Thr Asp Gly Tyr Ser Asp Ser Glu His
 2020 2025 2030
 Tyr Leu Pro Met Glu Gly Gln Thr Arg Ala Ala Ser Met Pro Arg Leu
 2035 2040 2045
 Pro Ala Glu Asn Gln Arg Arg Arg Gly Arg Pro Arg Gly Asn Asp Leu
 2050 2055 2060
 Ser Thr Ile Ser Asp Thr Ser Pro Met Lys Arg Ser Ala Ser Val Leu
 2065 2070 2075 2080
 Gly Pro Lys Ala Arg Arg Leu Asp Asp Tyr Ser Leu Glu Arg Val Pro
 2085 2090 2095
 Pro Glu Glu Asn Gln Arg Tyr His Gln Arg Arg Arg Asp Arg Gly His
 2100 2105 2110
 Arg Thr Ser Glu Arg Ser Leu Gly Arg Tyr Thr Asp Val Asp Thr Gly
 2115 2120 2125
 Leu Gly Thr Asp Leu Ser Met Thr Thr Gln Ser Gly Asp Leu Pro Ser
 2130 2135 2140
 Lys Asp Arg Asp Gln Asp Arg Gly Arg Pro Lys Asp Arg Lys His Arg
 2145 2150 2155 2160
 Pro His His His His His His His His His Pro Pro Ala Pro Asp
 2165 2170 2175
 Arg Asp Arg Tyr Ala Gln Glu Arg Pro Asp Thr Gly Arg Ala Arg Ala
 2180 2185 2190
 Arg Glu Gln Arg Trp Ser Arg Ser Pro Ser Glu Gly Arg Glu His Thr
 2195 2200 2205
 Thr His Arg Gln Gly Ser Ser Ser Val Ser Gly Ser Pro Ala Pro Ser
 2210 2215 2220
 Thr Ser Gly Thr Ser Thr Pro Arg Arg Gly Arg Arg Gln Leu Pro Gln
 2225 2230 2235 2240
 Thr Pro Cys Thr Pro Arg Pro Leu Val Ser Tyr Ser Pro Ala Pro Arg
 2245 2250 2255
 Arg Pro Ala Ala Arg Arg Met Ala Gly Pro Ala Ala Pro Pro Gly Gly
 2260 2265 2270
 Ser Pro Arg Gly Cys Arg Arg Ala Pro Arg Trp Pro Ala His Ala Pro
 2275 2280 2285
 Glu Gly Pro Arg Pro Arg Gly Ala Asp Tyr Thr Glu Pro Asp Ser Pro
 2290 2295 2300
 Arg Glu Pro Pro Gly Gly Ala His Asp Pro Ala Pro Arg Ser Pro Arg
 2295 2310 2315 2320

-continued

Thr Pro Arg Ala Ala Gly Cys Ala Ser Pro Arg His Gly Arg Arg Leu
2325 2330 2335

Pro Asn Gly Tyr Tyr Ala Gly His Gly Ala Pro Arg Pro Arg Thr Ala
2340 2345 2350

Arg Arg Gly Ala His Asp Ala Tyr Ser Glu Ser Glu Asp Asp Trp Cys
2355 2360 2365

<210> SEQ ID NO 9

<211> LENGTH: 6950

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

ggccaaagg	caagatactt	ctctggaaa	tgttgctgct	gatgctgctt	tacaaagtca	60
tacaatgagt	gtttggttt	agaaaagattt	tcataacttaa	aagattttca	tcttggaaat	120
acatcaagtg	aaaattaaat	tctttggga	aacattttcc	ttctgtatata	ttataacttgt	180
aatgggcgac	atggcaaaca	actcagttgc	ttacagtgg	gtgaaaaact	ctttgaagga	240
agctaatcat	gatggagact	ttgaaattac	gtcgcagag	ctgcgggctc	tcatggagct	300
caggtccaca	gatgcattac	aaaaataaca	ggaaagctat	ggagatgtct	atggaaatttgc	360
caccaaattt	aaaacatctc	ccaatgaagg	ttaagtgg	aaccctgcag	attttagaaag	420
aagagaagca	gtgtttggaa	agaattttat	acctcctaaa	aagccaaaaa	cctttttca	480
attagtatgg	gaagcattac	aagatgtcac	ttaattata	ttagaaatttgc	cagccatagt	540
atcattggc	ctttttttt	atcagcctcc	agaaggggat	aatgcacttt	gtggagaagt	600
ttctgttgg	gaggaagaag	gtgaagggtga	aactggttgg	attgaaggag	ctgcaatcct	660
cttgcgtgt	gtgtgtgtgg	tgttagtaac	agctttcaat	gactggagta	aggaaaaaca	720
gttttagaggt	ttgcagagcc	gaattgaaca	agaacagaag	ttcactgtca	tcaggggtgg	780
tcaggtcatt	cagatacctg	tagctgacat	tactgttgg	gatattgctc	aagtgaaata	840
tggtgatctt	cttccagctg	acggcatact	tattcaaggc	aacgatctta	aaattgtat	900
aaagtcattt	actgttgcata	cagatcatgt	taaaaaagtct	ttagataagg	atcccttact	960
tctatcagg	actcatgtaa	tggaggctc	tggagaatg	gtagttacag	ctgttaggtgt	1020
aaattctcaa	actggaaatta	tctttacctt	acttggagct	ggagggttgg	agggaaagagaa	1080
gaaagatgag	aaaaaaaagg	aaaagaaaaa	taagaaacaa	gatggagct	ttgagaatcg	1140
caacaaagca	aaagcccagg	atggcgcagc	catggaaatg	cagccatgt	agagtgaaga	1200
agggtggagat	ggtgatgaaa	aaataaaaaa	gaaagcaat	ttgccaaaaa	aggaaaaatc	1260
tgttttacaa	ggggaaactt	caaaactggc	tgttcagatt	ggcaaagcag	gtctgttgc	1320
gtctgccatc	acagttatca	ttcttagtatt	atattttgtc	attgacacct	tctgggttca	1380
gaaaagacca	tggcttgcgt	agtgcacacc	aatttatata	caataacttg	tgaagtctt	1440
cattatttgg	gttacagttt	tagtggcgc	agtgcgcagaa	ggtctccac	ttgcagtcac	1500
gatctcactg	gcttatttgc	tcaaaaaat	gatggaaat	aataacttag	taaggcatct	1560
ggatgcttgc	gaaaccatgg	gaaatgtac	agctatttgc	tcaaaaaaa	cagggacttt	1620
gacaatgaac	agaatgacag	tcgttcaagc	ttacataat	gaaaaacatt	ataaaaaggt	1680
tcctgaacca	gaagcttattc	caccaaataat	tttgccttat	tttgcacag	gaatttctgt	1740
gaatttgcgt	tatacatcaa	aaatatttgc	accagagaaa	gagggttgg	tacctcgta	1800

-continued

cgttggtaat	aaaactgaat	gtgccttgtt	gggacttctt	ttggattaa	aacgggatta	1860
tcaggatgtt	agaaatgaaa	taccagaaga	agcaactgtac	aaagtctaca	ccttcaattc	1920
tgttaggaag	tccatgagta	ctgtcctgaa	aaattcagat	ggaagttatc	aatattcag	1980
caagggtgca	tctgagataa	ttctgaaaaa	gtgtttcaaa	atcttgagtg	ctaatggtga	2040
ggcaaaagta	ttcagacaa	gggaccgtga	tgatattgta	aaaactgtga	ttgaacccgt	2100
ggcatcagaa	ggcttgagaa	ccatatgtct	tgcattcaga	gatttccag	caggagaacc	2160
agaaccagag	tgggataatg	aaaatgatat	tgtcacccgc	cttacatgca	ttgctgttgt	2220
ggggattgaa	gatcctgtga	gacctgaggt	gccagatgca	attaaaaagt	gtcagagggc	2280
tggaaattact	gtgcggatgg	tcactggtga	taatattaat	actgctcggg	ccattgtcac	2340
caaatgtggt	atttacatc	ctgggagaaga	ttttctgtgc	ctagaaggtta	aagatttaa	2400
cagaagaata	cgaaatgaaa	aaggagagat	tgagcaagag	aggatagaca	agatttggcc	2460
aaaacttcga	gtacttgcaa	gatcatctcc	tactgataag	catacactgg	ttaaaggttat	2520
aattgacagc	actgtctcag	accaacgcca	ggtttagtgc	gtaactggtg	atggtacaaa	2580
tgtatggcca	gcactaaaga	aagcagatgt	tggatttgca	atgggtatttgc	ctggactgta	2640
tgtagctaaa	gaagcatccg	atattattct	cacagatgac	aactttacaa	gcattgttaa	2700
acgcaggat	tggggaccaa	atgtctatga	cagcatctca	aaattccctc	agttcoact	2760
tactgttaat	gtagtagcag	tgattgttgc	ttttacgggc	gcctgcatttca	ctcaagactc	2820
accgcctaag	gctgtgcaga	tgctgtgggt	aaacctcata	atggatacac	tcgcctccct	2880
ggctctggca	acggaaccac	ccactgagtc	tctcttgctt	cgaaacctt	atggtagaaa	2940
taagcctctc	atctcacgta	caatgatgaa	aatattttgc	ggtcatgcat	tctataact	3000
tgttagtagtc	tttacactct	tatgtctgg	agaaaaagttt	tttgacatttgc	atagtggaaag	3060
aaatgctcct	ttgcatgetc	ctccctcaga	acattatact	attgttttta	atacctttgt	3120
gctgatgcaa	ctttcaacg	aaataaatgc	ccggaaaatt	catggtaaa	gaaatgtatt	3180
cgaaggaaatc	ttaacaatg	ccatcttctg	cacaatttgc	tttaggcactt	ttgtggtaca	3240
gataataatt	gtgcagtttgc	gtggaaaacc	tttcagttgt	tcagaactttt	caatagaaca	3300
gtggctatgg	tcaatattcc	taggaatggg	aacattactc	tggggccagc	ttatttcaac	3360
aattccaaatc	agccgtttaa	aattcctcaa	agaagctggt	catggAACAC	aaaaggaaga	3420
aataacccat	gaggaatttgc	cagaggatgt	tgaagagatt	gatcacgtc	aaagggagtt	3480
gcggcgtggc	caaatcttgt	ggttttaggg	tctgaacaga	atccaaacac	agatggatgt	3540
agtgaatgt	ttccagagtg	gaagttccat	tcaaggggct	ctaaggccgc	aaaccctccat	3600
cggccagccag	catcatgatg	taacaaatat	ttctaccctt	acacatgtac	tgtttccctc	3660
ttctactgt	tctactactg	tgggttatttgc	gagtggtgaa	tgcatttcgt	agttctttat	3720
atgaagggtt	agaaaaacccg	gaatcaagaa	gttcgattca	caactttatg	acacatcctg	3780
agtttaggt	agaagattca	gagcctcata	tcccccttat	tgtacact	gatgccgaaag	3840
atgatgctcc	tacaaaacgt	aactccagtc	ctccaccctc	tcccaacaaa	aataacaatg	3900
ctgttgcac	tggaaattc	cttacaatag	aaatgaacaa	gtctgctacc	tcttcattccc	3960
caggaagccc	actacatagt	ttggaaacat	cactctgatt	gtaagctgaa	tgttaacacaca	4020
ctagctgcac	tgtaaagaaa	caaattgaaa	ctgggtcttt	tcacatatttgc	tgatggacaa	4080

-continued

gctagtattc ttgtctttgg acttcaacag aagacacact tgcgtacgaatg tagatttatt	4140
tttttaaaaa aaaaaaaagca aagctttctg ccagactgag ggtgctttt ggggggtggg	4200
agaaatgaac tgacagataa acagtaactc agcgtaaatg acctgtggat catcgtatgt	4260
acccaagcat ggtcctgaac agtgtttag cagagcttta gtttatctca gtcctggatg	4320
ggagagagag ggaggagtaa aactgggcaa gacaaagagc cctggatcg tgaattgata	4380
gtttaatttc tgctgttgc ttttataat ttggatttatt tatgtttata aatgatacag	4440
atctgtttac aaggttttaa gatactttt ttgttccctgt tcatacatgg gaagcttcct	4500
tataactgtat gcagagaaaa attagtctt caaatactgc tgcgtatccaa ggaaaaaaa	4560
aaaaagggtt gtttctattt aactgtact aaattttgc ctacaattta ctttggtaaa	4620
tattgttagat aaattgtctaa tactaatagc caaatagata acactaatgc ttgtttaaa	4680
aaaacgaaaa aaacaaaaca aaacaaaaaa acatgagcag tgggtttcta atgaagttat	4740
gacatctgtc ctaatttagtt tcttaatatac atttactact taatgggtttt gaagcaactg	4800
ttaattttt aaaatttttgg aaaaacttgc tataaccctt tggcagaat tttagtagatt	4860
gtttaatgca ggacatcaac tacataatga aagatgcttgc aaatgctttt gtttccatgg	4920
gcaaatcaaa ttaaatcaga ctatcaaact acaagagaac cttagtcctt ttgttttg	4980
tctaaacatg ttatgtttagt gatgtcttgg tggactgtga gtggactgtta ttgttttc	5040
taatctttag agaccttcac acagcatgag ggctgttggc atttggaaaa aggttaataa	5100
gtagaagcat acatgttttc cttttgtttt tggaaacttgc ttgtaaacat aaataaaat	5160
gcccttttat taaataaattt catggcaatg tttttttaaa agacatttca gtttccatgg	5220
tcagttctgc atatacagtgc catcgtctgc ttctctttga ccacaattgg aatacatatc	5280
aagatggtaa cagccccatg aagtaccgc ttaattccctt gctggtaatc taatagaatc	5340
agattaagaa aattaacaca attttaccat ttgacataact cttcattgaa tttagttgt	5400
ctaaaatagt ctcacaggta aaaaacataa caacaatgac caccattgtc ttgtatacat	5460
ttattttaat tcttaatgtgg ccaataaca gattaggtaa aatgagatta catggaaaagg	5520
ttctaaaatg atatatttag atgcaataat taaccatatg ctggatgcc cacctataag	5580
ccaaattggat tcttttagtc tcttttctt tccttcatct accttcttctt ctttttgc	5640
tctgtcttta cccttttagaa actggcatgt gtttagggccc agatttacc agatacaca	5700
tcctcaggct gacataaaatgt ttacgaggac tgctactata actgagatgt ctgaccaagc	5760
aaagaaacat ttatgttgc tgcgttgcata tttttggggggaa tcttttctc	5820
accatcaaataat cttgcgttgcata atcctctgcata ggtcttttagaa acacattggc tctgtcagca	5880
tttggtaatttgc tacatttttgc catcaggat tttatgttgcata tcataactgtt aaaaagaggt	5940
agcatgatataat gctggaaagc atatttagact aggagtcaag acctgggttc tagtcttgc	6000
tttgcatttgcata atgtgttgcata tgcgttgcata tttttttttt tggttatgt	6060
actatcttgcata agaaccctt caactctaaa aattggaccat atcaactggc gagccacaa	6120
ggcataggaa atagagaaaa gggaaatgtt ttcagtttc tttttttttt tggttatgt	6180
taaatatgaa tgcgttgcata aaaaagatgtt gaaatgttgcata tgcgttgcata gatgttgc	6240
gagcttttgcata agtatttgcata tttcaactgcata aggcaatggc tttctacacc aactctaaatc	6300
cataaacggg tcttatgcata tctatgttgcata agtgcata gatgttgcata tgcgttgcata	6360

-continued

```

ctctttgaga cactgttaatt tctaccagaa atttccagag cattatgttag gtagaaaaaa 6420
atgcaggcaa gctgtttaag atcttggatc ccattatata gtatgtatag ctgaaatctg 6480
taattcaatc actttttctc ttttatcctc taaccaaaaa attgttaat tttgcattccc 6540
aatatgtttt aatctttgtt tatttttaa aaatcctttt ctcctcatca ttgcctttt 6600
tgtgggtgtt aatagactta ctgcacttt gaagatgagt tactccttg catcttacaa 6660
atatgtgtata tggttaattt cataacagat gtcagtttg aaccaagaat tggtgatttg 6720
tttataagaa aaaaactggc ttcatttctg tgaaattgct ctggaaaat ttcttttac 6780
acgtgtaaac caactgagat accgtgtatgg tggatttc ttcaatgtat gcttaccatc 6840
tatttttagcc actgagcctt ttatttttg tctatttgta aagtttattt gtcttaactc 6900
attnaataaaa tatactgttt atctgtttct gaaaaaaaaaaaaaaaaaaaa 6950

```

<210> SEQ ID NO 10

<211> LENGTH: 1176

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

Met	Gly	Asp	Met	Ala	Asn	Asn	Ser	Val	Ala	Tyr	Ser	Gly	Val	Lys	Asn
1				5				10				15			

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

Glu	Leu	Arg	Ala	Leu	Met	Glu	Leu	Arg	Ser	Thr	Asp	Ala	Leu	Arg	Lys
				35			40				45				

Ile	Gln	Glu	Ser	Tyr	Gly	Asp	Val	Tyr	Gly	Ile	Cys	Thr	Lys	Leu	Lys
				50			55			60					

Thr	Ser	Pro	Asn	Glu	Gly	Leu	Ser	Gly	Asn	Pro	Ala	Asp	Leu	Glu	Arg
65				70				75		80					

Arg	Glu	Ala	Val	Phe	Gly	Lys	Asn	Phe	Ile	Pro	Pro	Lys	Lys	Pro	Lys
				85			90		95						

Thr	Phe	Leu	Gln	Leu	Val	Trp	Glu	Ala	Leu	Gln	Asp	Val	Thr	Leu	Ile
				100			105			110					

Ile	Leu	Glu	Ile	Ala	Ala	Ile	Val	Ser	Leu	Gly	Leu	Ser	Phe	Tyr	Gln
115				120				125							

Pro	Pro	Glu	Gly	Asp	Asn	Ala	Leu	Cys	Gly	Glu	Val	Ser	Val	Gly	Glu
130				135				140							

Glu	Glu	Gly	Glu	Thr	Gly	Trp	Ile	Glu	Gly	Ala	Ala	Ile	Leu		
145				150			155			160					

Leu	Ser	Val	Val	Cys	Val	Val	Leu	Val	Thr	Ala	Phe	Asn	Asp	Trp	Ser
				165			170		175						

Lys	Glu	Lys	Gln	Phe	Arg	Gly	Leu	Gln	Ser	Arg	Ile	Glu	Gln	Glu	Gln
				180			185		190						

Lys	Phe	Thr	Val	Ile	Arg	Gly	Gly	Gln	Val	Ile	Gln	Ile	Pro	Val	Ala
195				200				205							

Asp	Ile	Thr	Val	Gly	Asp	Ile	Ala	Gln	Val	Lys	Tyr	Gly	Asp	Leu	Leu
210				215			220								

Pro	Ala	Asp	Gly	Ile	Leu	Ile	Gln	Gly	Asn	Asp	Leu	Lys	Ile	Asp	Glu
225				230			235		240						

Ser	Ser	Leu	Thr	Gly	Glu	Ser	Asp	His	Val	Lys	Lys	Ser	Leu	Asp	Lys
245				250			255			255					

-continued

Asp Pro Leu Leu Leu Ser Gly Thr His Val Met Glu Gly Ser Gly Arg
 260 265 270
 Met Val Val Thr Ala Val Gly Val Asn Ser Gln Thr Gly Ile Ile Phe
 275 280 285
 Thr Leu Leu Gly Ala Gly Gly Glu Glu Glu Lys Lys Asp Glu Lys
 290 295 300
 Lys Lys Glu Lys Lys Asn Lys Lys Gln Asp Gly Ala Ile Glu Asn Arg
 305 310 315 320
 Asn Lys Ala Lys Ala Gln Asp Gly Ala Ala Met Glu Met Gln Pro Leu
 325 330 335
 Lys Ser Glu Glu Gly Gly Asp Gly Asp Glu Lys Asp Lys Lys Lys Ala
 340 345 350
 Asn Leu Pro Lys Lys Glu Lys Ser Val Leu Gln Gly Lys Leu Thr Lys
 355 360 365
 Leu Ala Val Gln Ile Gly Lys Ala Gly Leu Leu Met Ser Ala Ile Thr
 370 375 380
 Val Ile Ile Leu Val Leu Tyr Phe Val Ile Asp Thr Phe Trp Val Gln
 385 390 395 400
 Lys Arg Pro Trp Leu Ala Glu Cys Thr Pro Ile Tyr Ile Gln Tyr Phe
 405 410 415
 Val Lys Phe Phe Ile Ile Gly Val Thr Val Leu Val Val Ala Val Pro
 420 425 430
 Glu Gly Leu Pro Leu Ala Val Thr Ile Ser Leu Ala Tyr Ser Val Lys
 435 440 445
 Lys Met Met Lys Asp Asn Asn Leu Val Arg His Leu Asp Ala Cys Glu
 450 455 460
 Thr Met Gly Asn Ala Thr Ala Ile Cys Ser Asp Lys Thr Gly Thr Leu
 465 470 475 480
 Thr Met Asn Arg Met Thr Val Val Gln Ala Tyr Ile Asn Glu Lys His
 485 490 495
 Tyr Lys Lys Val Pro Glu Pro Glu Ala Ile Pro Pro Asn Ile Leu Ser
 500 505 510
 Tyr Leu Val Thr Gly Ile Ser Val Asn Cys Ala Tyr Thr Ser Lys Ile
 515 520 525
 Leu Pro Pro Glu Lys Glu Gly Leu Pro Arg His Val Gly Asn Lys
 530 535 540
 Thr Glu Cys Ala Leu Leu Gly Leu Leu Leu Asp Leu Lys Arg Asp Tyr
 545 550 555 560
 Gln Asp Val Arg Asn Glu Ile Pro Glu Glu Ala Leu Tyr Lys Val Tyr
 565 570 575
 Thr Phe Asn Ser Val Arg Lys Ser Met Ser Thr Val Leu Lys Asn Ser
 580 585 590
 Asp Gly Ser Tyr Arg Ile Phe Ser Lys Gly Ala Ser Glu Ile Ile Leu
 595 600 605
 Lys Lys Cys Phe Lys Ile Leu Ser Ala Asn Gly Glu Ala Lys Val Phe
 610 615 620
 Arg Pro Arg Asp Arg Asp Asp Ile Val Lys Thr Val Ile Glu Pro Met
 625 630 635 640
 Ala Ser Glu Gly Leu Arg Thr Ile Cys Leu Ala Phe Arg Asp Phe Pro
 645 650 655

-continued

Ala Gly Glu Pro Glu Pro Glu Trp Asp Asn Glu Asn Asp Ile Val Thr
 660 665 670
 Gly Leu Thr Cys Ile Ala Val Val Gly Ile Glu Asp Pro Val Arg Pro
 675 680 685
 Glu Val Pro Asp Ala Ile Lys Lys Cys Gln Arg Ala Gly Ile Thr Val
 690 695 700
 Arg Met Val Thr Gly Asp Asn Ile Asn Thr Ala Arg Ala Ile Ala Thr
 705 710 715 720
 Lys Cys Gly Ile Leu His Pro Gly Glu Asp Phe Leu Cys Leu Glu Gly
 725 730 735
 Lys Asp Phe Asn Arg Arg Ile Arg Asn Glu Lys Gly Glu Ile Glu Gln
 740 745 750
 Glu Arg Ile Asp Lys Ile Trp Pro Lys Leu Arg Val Leu Ala Arg Ser
 755 760 765
 Ser Pro Thr Asp Lys His Thr Leu Val Lys Gly Ile Ile Asp Ser Thr
 770 775 780
 Val Ser Asp Gln Arg Gln Val Val Ala Val Thr Gly Asp Gly Thr Asn
 785 790 795 800
 Asp Gly Pro Ala Leu Lys Lys Ala Asp Val Gly Phe Ala Met Gly Ile
 805 810 815
 Ala Gly Thr Asp Val Ala Lys Glu Ala Ser Asp Ile Ile Leu Thr Asp
 820 825 830
 Asp Asn Phe Thr Ser Ile Val Lys Ala Val Met Trp Gly Arg Asn Val
 835 840 845
 Tyr Asp Ser Ile Ser Lys Phe Leu Gln Phe Gln Leu Thr Val Asn Val
 850 855 860
 Val Ala Val Ile Val Ala Phe Thr Gly Ala Cys Ile Thr Gln Asp Ser
 865 870 875 880
 Pro Leu Lys Ala Val Gln Met Leu Trp Val Asn Leu Ile Met Asp Thr
 885 890 895
 Leu Ala Ser Leu Ala Leu Ala Thr Glu Pro Pro Thr Glu Ser Leu Leu
 900 905 910
 Leu Arg Lys Pro Tyr Gly Arg Asn Lys Pro Leu Ile Ser Arg Thr Met
 915 920 925
 Met Lys Asn Ile Leu Gly His Ala Phe Tyr Gln Leu Val Val Phe
 930 935 940
 Thr Leu Leu Phe Ala Gly Glu Lys Phe Phe Asp Ile Asp Ser Gly Arg
 945 950 955 960
 Asn Ala Pro Leu His Ala Pro Pro Ser Glu His Tyr Thr Ile Val Phe
 965 970 975
 Asn Thr Phe Val Leu Met Gln Leu Phe Asn Glu Ile Asn Ala Arg Lys
 980 985 990
 Ile His Gly Glu Arg Asn Val Phe Glu Gly Ile Phe Asn Asn Ala Ile
 995 1000 1005
 Phe Cys Thr Ile Val Leu Gly Thr Phe Val Val Gln Ile Ile Ile Val
 1010 1015 1020
 Gln Phe Gly Gly Lys Pro Phe Ser Cys Ser Glu Leu Ser Ile Glu Gln
 1025 1030 1035 1040
 Trp Leu Trp Ser Ile Phe Leu Gly Met Gly Thr Leu Leu Trp Gly Gln
 1045 1050 1055
 Leu Ile Ser Thr Ile Pro Thr Ser Arg Leu Lys Phe Leu Lys Glu Ala

-continued

1060	1065	1070	
Gly His Gly Thr Gln Lys Glu Glu Ile Pro Glu Glu Glu Leu Ala Glu			
1075	1080	1085	
Asp Val Glu Glu Ile Asp His Ala Glu Arg Glu Leu Arg Arg Gly Gln			
1090	1095	1100	
Ile Leu Trp Phe Arg Gly Leu Asn Arg Ile Gln Thr Gln Met Asp Val			
1105	1110	1115	1120
Val Asn Ala Phe Gln Ser Ser Ile Gln Gly Ala Leu Arg Arg			
1125	1130	1135	
Gln Pro Ser Ile Ala Ser Gln His His Asp Val Thr Asn Ile Ser Thr			
1140	1145	1150	
Pro Thr His Val Val Phe Ser Ser Ser Thr Ala Ser Thr Thr Val Gly			
1155	1160	1165	
Tyr Ser Ser Gly Glu Cys Ile Ser			
1170	1175		

<210> SEQ ID NO 11

<211> LENGTH: 2670

<212> TYPE: DNA

<213> ORGANISM: Xenopus tropicalis

<400> SEQUENCE: 11

ccccgggctt	gagggtgcat	tatcataaaag	acttgttagtt	cttttgcag	attttttag	60
ccatacatga	ctgggtgtag	tgagatgtct	acctgcctgc	atcttcaata	cacttgagag	120
tttctaagtgc	ctcataatac	ctccgttagct	ttcatcgatt	ttatttctgt	tagctgtct	180
ataaaaagtgc	ctttttctct	ttgctttta	tttattgttag	atgggatggt	cttttctt	240
gttctctctc	tctcttgcgt	agoaagctgt	cacaatctct	gattccttgc	agatggatgt	300
agtgaatgtct	ttccaaagtgc	gaagttccat	tcagggggct	ttaaggccgc	aaccctccat	360
cgccagccag	catcatgtat	taacaaatat	ttctaccctt	acacatgtat	tgttttcttc	420
ctctactgtct	tctactactg	tgggtgtgat	gtgtgtgttc	ctaagtgcata	gaaattaaca	480
catacctaatt	cacataccta	acgtttctca	ttttttcttc	agtaatcaat	ctcattgttc	540
aggctgtctca	taaagcttcc	cttcacaag	gataaaaagt	cctgtgttgt	tttatatttt	600
ttttccatag	aaggctatttgc	aactgggttgg	aatggtagca	aatgcaagtt	ttgaaaaatgt	660
tctgcaaccc	tgcacgcctt	tatattataag	tatattgtct	cgtttatgtt	tatatttggaa	720
agttgattgt	gctgctgtttaa	attttcgagt	atataccgtt	ttttttatgt	cagaaggcat	780
gatgtttggc	tttgggata	ttttacactt	atatatctt	gtttcacat	ttctatatttt	840
tgttaggcaac	cacataaaagg	ataaattctt	tacagtctt	ttaaattttc	ttgtgtactac	900
taggatatgg	actagctaac	tattgacata	gttgggttgc	aatgtgacca	tgggtggca	960
gaagacatct	agacaaatac	tctatccctt	ttctctaaaa	gtggaaatct	atgctcagag	1020
agagatcctt	cccaattgcc	ttccctacat	tctctccatg	accatgttcc	gctgtgacat	1080
aatcccaggc	agtttagctt	tatggcttat	gaagtggatg	tgatttttgt	taggttgtgt	1140
gtttttggta	gcaaagggtcc	catttatgag	aactaaatcc	acatgaggca	aaacttgc	1200
tagaaaaagg	gttttgcataa	ggaaagctga	tatttgcctt	ccttagcaat	gctctccagc	1260
aagtgtttgc	cctaaagaat	ctccttctgt	gtgcttattac	tattagaact	taatataatg	1320
taatcgaaaa	tttgaacagc	gttgctatac	aagaaagcag	agcacgaaata	acactttatc	1380

-continued

ttgggtttac tatccacaaa aatgtctatt tctatccaca gaattacttc tactacttca 1440
gctgttgttgc aactagtgta ttggccaaaa ttaattttat ataccgtatg ttttatgttg 1500
gcattgtgg ttgtgagtgaa ggcatgtgga ttctagtaaa caggatggca acctgtatagg 1560
aacatataaa accactctaa gaacagtca tgacccgggt tgcacccgtc tcaaaacata 1620
taaagttca aggtttcca acatgtatag tttaataaa tatatccata tggcctacaa 1680
attcataggg agtacttcatttactacatgt ttgtgagcta catctaccac tgaagccccct 1740
gatctacttg tatatacgctg taatgtgatt ttgttcatgc aagaaaaata tgcacataga 1800
tcactttctat gcttgctaat tggtccgcga ttgtatcaat tagtcttaac attttggcac 1860
cagacaatgt tatatggagg ttttagttt ctatgggtct ctcatacaca aacccttctt 1920
cttttgggg gaaaacctgc tgcgtggcag atattgggtc tggtaagcag ctggaaag 1980
ctgagtaata ccagtgacaaag tggtgagttt tctacctgat ttattaaggg caccaccta 2040
atgcttagata aaaataccaa tggtttgcata cccattataa attgcggcaa gacatgttgt 2100
ttaggttttc cagggtcagca gattaacata gaataagaca aaaacaata tcaaggcctg 2160
ttgctgcag ttatcttaca cacttgctct tatctattct tctctaaatgta ctattcttgt 2220
gttgtatact tgtagttagca gtctctctct ctatgggtg gataacaaca atccaaatgc 2280
aaattttggta tttaaagggtt gaccacaatt gttttttttt atttgaagat tgaactgttgt 2340
ttaaacaaat gttttttttt ccccttaggg ttccagatta tttaaaaat gaaaaaaagtgc 2400
atgtttaaca atgtactttt acatttccaa gggttataac ttttgcattcc agcatcaagt 2460
atatttaatt aaacttctat atagtcttattt cttgcacaaa aatgcactta gtaatcatgt 2520
ctgttagcgc tcacattagg gatgcattaa acctatcggtt tttggattcg ggcaaatacc 2580
aaatcttctat cacttagattt gtttgagtttac ttaactgaac ctttttttc gtattcaagt 2640
ttgtgtataat caaaaaaaaaaaaaaa 2670

<210> SEQ ID NO 12
<211> LENGTH: 59
<212> TYPE: PRT
<213> ORGANISM: *Xenopus tropicalis*

<400> SEQUENCE: 12

```

Met Asp Val Val Asn Ala Phe Gln Ser Gly Ser Ser Ile Gln Gly Ala
          5                   10                  15

```

Leu Arg Arg Gln Pro Ser Ile Ala Ser Gln His His Asp Val Thr Asn
20 25 30

Ile Ser Thr Pro Thr His Val Val Phe Ser Ser Ser Thr Ala Ser Thr
35 40 45

Thr Val Gly Cys Glu Cys Val Phe Leu Ser Ala
50 55

<210> SEQ ID NO 13
<211> LENGTH: 3250
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

```
ccggaggatt ttgaggacac ttgtggagag ctgcgaattcc agaatgtga aatttgttagg 60  
ttgtgacagt tggaagtgtc atgtacaaca tgcgccgatt aagtctttca cccaccttt 120
```

-continued

caatgggg tcatctgtta gttactgtga gtcttatt ttccatgtg gaccatgtaa 180
ttgctgagac agaaatggaa ggagaaggaa atgaaactgg tgaatgtact ggatcatatt 240
actgtaaagaa aggggtatt ttgcccattt gggAACCCCA agacccttct tttggggaca 300
aaattgctag agctactgtg tattttgtgg ccatggtcta catgtttctt ggagtctcta 360
tcatactgtc tcgggtcatg tcctctatag aagtcatcac atctcaagaa aaagaaataa 420
ccataaaagaa acccaatggg gagaccacca agacaactgt gaggatctgg aatgaaacag 480
tttcttaacct gacccatgtg gcccggat cttctgtcc tgagattctc cttagttaa 540
ttgaagtgtg tggccataaac ttcaactgcag gagacctcg tcctagcacc atcggtggaa 600
gtgctgcatt caatatgttc atcattattt cactctgtgt ttatgtggg cctgacggag 660
agacaaggaa gattaagcat ttgcgtgtct tctttgtgc acgcagctgg acatctttg 720
cctacacccgt gctttacatt atttgtctg tcatacttcc tgggtgtgt gagggtctggg 780
aagggttgct tactttctt ttctttccca tctgtgttgtt gttcgcttgg gtacggata 840
ggagacttct gttttacaag tatgtctaca agaggtatcg agctggcaag cagaggggaa 900
tgattattgtt acatgaaggaa gacaggccat cttctaaagac tgaaatgtt atggacgggaa 960
aagtggtcaa ttctcatgtt gaaaatttct tagatggtgc tctgggtctg gaggtggatg 1020
agagggaccat agatgtgaa gaagcttaggc gagaaatggc taggattctt aaggaactta 1080
agcagaagca tccagataaa gaaatagac aattaataga attagctaact taccaagtcc 1140
taagtgcacca gaaaaaaagt agagcatttt atcgcattca agctactcgc ctcatgactg 1200
gagctggcaa cattttaaag aggcatgcag ctgaccaagc aaggaaggct gtcagcatgc 1260
acgaggtcaa cactgaagtg actgaaaatg accctgttag taagatctt tttgaacaag 1320
ggacatatca gtgtctggag aactgtggta ctgtggccct taccattatc cgccaggtg 1380
gtgatttgac taacactgtg tttgttgact tcagaacaga ggatggcaca gcaaattgtg 1440
ggtctgatta tgaattttact gaaggaactg tgggtttaa gcctgggtat acccagaagg 1500
aaatcagatg gggatcata gatgtatgata tctttgaggaa ggatggaaaat ttcttggc 1560
atctcagcaa tgtcaaagta tcttctgaag cttcagaaga tggcatactg gaagccaatc 1620
atgtttctac acttgcgttgc ctggatctc cttccactgc cactgttaact atttttgtat 1680
atgaccacgc aggcatttt acttttgagg aacctgtgc tcatgtgagt gagagcattt 1740
gcatcatgga ggtgaaagat ttgagaacat ctggagctcg agggaaatgtt atcggtccat 1800
ataaaaaccat cgaagggact gccagaggtg gaggggaggaa ttttggggacttgg 1860
agctcgaatt ccagaatgtat gaaattgtca aaacaatatac agtcaaggta attgtatgt 1920
aggaggtatgtt gaaaaacaag accttcttcc ttgagattgg agagccccgc ctgggtggaga 1980
tgagtgtgaa gaaagccctg ttattgtatg agcttgggtt gttcacaata acaggaaaat 2040
acctgtttgg ccaacactgtc ttcatggagg ttcatgtctg agaacatccg attctctcta 2100
ctgtatcac cattgcagac gaatatgtat acaaggcagcc actgaccacgc aaagagggaa 2160
aggagggcgc cattgcagaa atggggcgcc ccatccctggg agagcacacc aagttggaa 2220
tgcattgtt gaaatcctat gaattcaaga gtactgtggaa caaacttatt aagaagacaa 2280
acctggccct tttgggttggg actaacatgtt ggagagaaca gttcattgtt gctatctactg 2340
tcagtgctgg ggaatgtatgat gacgacgtat aatgtggggaa agagaagctg cccctctgtt 2400

-continued

tcgattacgt gatgcacttt ctgactgtgt tctggaaggt cctgtttgcc ttcgcccccc	2460
ctactgaata ctggaatggc tgggcgtgtt tcattgtctc catcctcatg attggcctac	2520
tgacagctt cattggagac ctggctccc actttggctg caccatggc ctgaaagatt	2580
ctgtgactgc agtcgtgttc gtcgcacttg gaacatca gtccagacaca ttggccagca	2640
aagtggcaggc caccaggac cagtatgcag acgcctccat aggttaacgtc acggggagca	2700
acgcgggtgaa tgtcttcctg ggaatcggtg tggcctggtc catcgctgcc atctaccacg	2760
cagccaatgg ggaacagttc aaagtgtccc ctggcacact agctttctct gtcactctct	2820
tcaccattt tgcttcatac aatgtggggg tgctgtgtt tcggcggagg ccagaaatcg	2880
gaggtgagct ggggtggggcc cgactgcca agtcctcac atcctgcctc tttgtgtcc	2940
tatggcttctt gtacattttc ttctccccc tggaggccca ctggcacata aaaggcttct	3000
aaaaggaacta tcagatatacg taaattata tatatacata tatatacata aaaattatgt	3060
ataatggaca gaggaaactg acatttgca tgttcaactt cctgctgatg gaatccagct	3120
tcaagagcat actctgtact agggccgaag taaaaaacca tcaccccca ttcccagggg	3180
catcatcatg ttcaacaagg catggaggca gggccatctt tgcagctcag tctagaaggg	3240
ctgcactctc	3250

<210> SEQ_ID NO 14

<211> LENGTH: 973

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

Met Tyr Asn Met Arg Arg Leu Ser Leu Ser Pro Thr Phe Ser Met Gly			
1	5	10	15
Phe His Leu Leu Val Thr Val Ser Leu Leu Phe Ser His Val Asp His			
20	25	30	
Val Ile Ala Glu Thr Glu Met Glu Gly Glu Gly Asn Glu Thr Gly Glu			
35	40	45	
Cys Thr Gly Ser Tyr Tyr Cys Lys Lys Gly Val Ile Leu Pro Ile Trp			
50	55	60	
Glu Pro Gln Asp Pro Ser Phe Gly Asp Lys Ile Ala Arg Ala Thr Val			
65	70	75	80
Tyr Phe Val Ala Met Val Tyr Met Phe Leu Gly Val Ser Ile Ile Ala			
85	90	95	
Asp Arg Phe Met Ser Ser Ile Glu Val Ile Thr Ser Gln Glu Lys Glu			
100	105	110	
Ile Thr Ile Lys Lys Pro Asn Gly Glu Thr Thr Lys Thr Thr Val Arg			
115	120	125	
Ile Trp Asn Glu Thr Val Ser Asn Leu Thr Leu Met Ala Leu Gly Ser			
130	135	140	
Ser Ala Pro Glu Ile Leu Leu Ser Val Ile Glu Val Cys Gly His Asn			
145	150	155	160
Phe Thr Ala Gly Asp Leu Gly Pro Ser Thr Ile Val Gly Ser Ala Ala			
165	170	175	
Phe Asn Met Phe Ile Ile Ala Leu Cys Val Tyr Val Val Pro Asp			
180	185	190	
Gly Glu Thr Arg Lys Ile Lys His Leu Arg Val Phe Phe Val Thr Ala			

-continued

195	200	205
Ala Trp Ser Ile Phe Ala Tyr Thr Trp Leu Tyr Ile Ile Leu Ser Val		
210	215	220
Ile Ser Pro Gly Val Val Glu Val Trp Glu Gly Leu Leu Thr Phe Phe		
225	230	235
Phe Phe Pro Ile Cys Val Val Phe Ala Trp Val Ala Asp Arg Arg Leu		
245	250	255
Leu Phe Tyr Lys Tyr Val Tyr Lys Arg Tyr Arg Ala Gly Lys Gln Arg		
260	265	270
Gly Met Ile Ile Glu His Glu Gly Asp Arg Pro Ser Ser Lys Thr Glu		
275	280	285
Ile Glu Met Asp Gly Lys Val Val Asn Ser His Val Glu Asn Phe Leu		
290	295	300
Asp Gly Ala Leu Val Leu Glu Val Asp Glu Arg Asp Gln Asp Asp Glu		
305	310	315
Glu Ala Arg Arg Glu Met Ala Arg Ile Leu Lys Glu Leu Lys Gln Lys		
325	330	335
His Pro Asp Lys Glu Ile Glu Gln Leu Ile Glu Leu Ala Asn Tyr Gln		
340	345	350
Val Leu Ser Gln Gln Lys Ser Arg Ala Phe Tyr Arg Ile Gln Ala		
355	360	365
Thr Arg Leu Met Thr Gly Ala Gly Asn Ile Leu Lys Arg His Ala Ala		
370	375	380
Asp Gln Ala Arg Lys Ala Val Ser Met His Glu Val Asn Thr Glu Val		
385	390	395
Thr Glu Asn Asp Pro Val Ser Lys Ile Phe Phe Glu Gln Gly Thr Tyr		
405	410	415
Gln Cys Leu Glu Asn Cys Gly Thr Val Ala Leu Thr Ile Ile Arg Arg		
420	425	430
Gly Gly Asp Leu Thr Asn Thr Val Phe Val Asp Phe Arg Thr Glu Asp		
435	440	445
Gly Thr Ala Asn Ala Gly Ser Asp Tyr Glu Phe Thr Glu Gly Thr Val		
450	455	460
Val Phe Lys Pro Gly Asp Thr Gln Lys Glu Ile Arg Val Gly Ile Ile		
465	470	475
Asp Asp Asp Ile Phe Glu Asp Glu Asn Phe Leu Val His Leu Ser		
485	490	495
Asn Val Lys Val Ser Ser Glu Ala Ser Glu Asp Gly Ile Leu Glu Ala		
500	505	510
Asn His Val Ser Thr Leu Ala Cys Leu Gly Ser Pro Ser Thr Ala Thr		
515	520	525
Val Thr Ile Phe Asp Asp His Ala Gly Ile Phe Thr Phe Glu Glu		
530	535	540
Pro Val Thr His Val Ser Glu Ser Ile Gly Ile Met Glu Val Lys Val		
545	550	555
Leu Arg Thr Ser Gly Ala Arg Gly Asn Val Ile Val Pro Tyr Lys Thr		
565	570	575
Ile Glu Gly Thr Ala Arg Gly Gly Glu Asp Phe Glu Asp Thr Cys		
580	585	590
Gly Glu Leu Glu Phe Gln Asn Asp Glu Ile Val Lys Thr Ile Ser Val		
595	600	605

-continued

Lys Val Ile Asp Asp Glu Glu Tyr Glu Lys Asn Lys Thr Phe Phe Leu
610 615 620

Glu Ile Gly Glu Pro Arg Leu Val Glu Met Ser Glu Lys Lys Ala Leu
625 630 635 640

Leu Leu Asn Glu Leu Gly Gly Phe Thr Ile Thr Gly Lys Tyr Leu Phe
645 650 655

Gly Gln Pro Val Phe Arg Lys Val His Ala Arg Glu His Pro Ile Leu
660 665 670

Ser Thr Val Ile Thr Ile Ala Asp Glu Tyr Asp Asp Lys Gln Pro Leu
675 680 685

Thr Ser Lys Glu Glu Glu Arg Arg Ile Ala Glu Met Gly Arg Pro
690 695 700

Ile Leu Gly Glu His Thr Lys Leu Glu Val Ile Ile Glu Glu Ser Tyr
705 710 715 720

Glu Phe Lys Ser Thr Val Asp Lys Leu Ile Lys Lys Thr Asn Leu Ala
725 730 735

Leu Val Val Gly Thr Asn Ser Trp Arg Glu Gln Phe Ile Glu Ala Ile
740 745 750

Thr Val Ser Ala Gly Glu Asp Asp Asp Asp Glu Cys Gly Glu Glu
755 760 765

Lys Leu Pro Ser Cys Phe Asp Tyr Val Met His Phe Leu Thr Val Phe
770 775 780

Trp Lys Val Leu Phe Ala Phe Val Pro Pro Thr Glu Tyr Trp Asn Gly
785 790 795 800

Trp Ala Cys Phe Ile Val Ser Ile Leu Met Ile Gly Leu Leu Thr Ala
805 810 815

Phe Ile Gly Asp Leu Ala Ser His Phe Gly Cys Thr Ile Gly Leu Lys
820 825 830

Asp Ser Val Thr Ala Val Val Phe Val Ala Leu Gly Thr Ser Val Pro
835 840 845

Asp Thr Phe Ala Ser Lys Val Ala Ala Thr Gln Asp Gln Tyr Ala Asp
850 855 860

Ala Ser Ile Gly Asn Val Thr Gly Ser Asn Ala Val Asn Val Phe Leu
865 870 875 880

Gly Ile Gly Val Ala Trp Ser Ile Ala Ala Ile Tyr His Ala Ala Asn
885 890 895

Gly Glu Gln Phe Lys Val Ser Pro Gly Thr Leu Ala Phe Ser Val Thr
900 905 910

Leu Phe Thr Ile Phe Ala Phe Ile Asn Val Gly Val Leu Leu Tyr Arg
915 920 925

Arg Arg Pro Glu Ile Gly Glu Leu Gly Gly Pro Arg Thr Ala Lys
930 935 940

Leu Leu Thr Ser Cys Leu Phe Val Leu Leu Trp Leu Leu Tyr Ile Phe
945 950 955 960

Phe Ser Ser Leu Glu Ala Tyr Cys His Ile Lys Gly Phe
965 970

<210> SEQ ID NO 15

<211> LENGTH: 3344

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

-continued

<400> SEQUENCE: 15

ggatcgggca	ggtcagccct	cagagctgg	cggtttctca	tccageccat	gcagacagag	60
acggcttgcac	agaggttgg	gcagttggaa	gtcttattgt	acaacatgt	tcgattaagt	120
ctccccaccca	atgtttcaat	gggatttgcgt	ctggtagctc	tggtggtct	cttgggttcc	180
catgttgcacc	atataactgc	agatacacag	gcagaaaacag	gaggaaatga	aaccactgaa	240
tgtactggct	catattactg	taagaaagg	gtgatcttc	ccatgggaa	accccaagac	300
ccatcttttgc	gggacaaaat	tgctagagca	actgtgtatt	tttgtggccat	ggtctacatg	360
ttccttggag	tttcttattat	tgagacccgg	tttatgtcct	ctatagaggt	catcacatct	420
caagagaaag	aaataacat	aaagaaaccg	aatggagaga	ccaccaagac	gacggtgaga	480
atctggaaacg	agactgtgtc	gaacctgacc	ttgatggccc	tgggatcttc	tgctcttgag	540
attttccctgt	cagtcattga	agtgtgcgg	cataacttca	ccgcagggg	cctgggtccc	600
agcaggatcg	tgggaagtgc	tgcctttaac	atgttcatca	taatcgact	ctgtgtttac	660
gtggtccctg	atggagagac	aaggaagatc	aagcatctgc	gtgtgttctt	tgtgacagca	720
gcctggagca	tctttgccta	tacctggctt	tatataatct	tgtctgtcag	ctctccctgg	780
gttgtggagg	tctgggaagg	cttgcttact	ttcttcttct	ttcccatctg	cgttgtgttc	840
gcgtgggtag	cagacaggcg	gcttctttt	tacaagtatg	tctacaacgc	gtacaggggcc	900
ggcaaggcaga	gggggatgt	cattgaacat	gaaggagaca	gaccagcttc	caaaactgaa	960
atcgaaatgg	atgggaaagt	ggtcaactct	catgttgaca	atttctttaga	tggggctctg	1020
gttttggaaag	ttgatgagag	ggaccaagat	gatgaggaag	ccaggcgtga	gatggcaagg	1080
attctgaagg	aacttaagca	gaagcatctc	gagaaagaaa	ttgagcaatt	aatagaattt	1140
gccaaactacc	aggccctaaag	tcaacacag	aaaagccgag	cattttacag	gattcaagct	1200
actcgccctga	tgaccggagc	tggcaacatc	ttgaagaggc	acgcagctga	tcaagcaagg	1260
aggcgtgtca	gtatgcgtca	agtcaacatg	gaaatggctg	aaaacgaccc	agtcagtaag	1320
atcttctttg	agcaagggac	ataccagtgt	ctagagaact	gtggtaactgt	ggccctcacc	1380
attatgcgca	gagggggcga	cttgagcacc	actgtgtttg	ttgacttcag	gacagaagac	1440
ggcacagcca	atgctgggtc	tgattatgaa	ttcacggaa	ggactgtgtat	cttcaaacc	1500
ggggagaccc	agaaggaaat	cagagttggc	atcattgtg	atgatatctt	tgaagaagat	1560
gaaaacttcc	ttgtgcattct	tagcaatgtc	agagtctttt	cagatgttcc	agaagatggc	1620
atactagaat	ccaatcacgc	ttcttcattt	gcttgtcttgc	ggtcacccag	cactgcccacc	1680
ataaccattt	ttgataatga	ccatacaggc	atctttacat	ttgaggaacc	cgtgactcac	1740
gtgagcgcaga	gcattggcat	catggaggtg	aaaggtttga	gaacctctgg	agctcgagga	1800
aatgttatca	ttcccttacaa	aactattgaa	ggcacagccc	gaggtggagg	ggaagacttt	1860
gaggacaccc	ttggagtgcc	ggaattccag	aatgatgaaa	tagtcaaac	aatatcgatc	1920
aaggtaataa	tcgatgacga	ggagttatgag	aaaaacaaga	ctttcattga	gattggagag	1980
ccccgtctgg	tggagatgag	tgagaagaaa	gccctgttgt	tgaatgagct	tggtggtttc	2040
acattaacag	gaaaagagat	gtatggccaa	cctatcttca	ggaagggtcca	tgctagagat	2100
cateccgatc	cctctactgt	aatcaccatc	tcagaggaat	atgatgacaa	gcagccactg	2160
accagcaaag	aagaccagga	gaggcgcatt	ggggaaatgg	ggcgccccat	cctaggcggag	2220

-continued

cacaccaagc tggagggtat catcgaaagag tcttacgaat tcaagagcac tgtggacaaa	2280
ctcattaaga agacgaaacct ggcccttgcgt gttggggacca acagctggag agagcgttt	2340
atcgaagcca tcactgttag cgctggggaa cgtgacgatg atgatgaata tggggaggag	2400
aagctgcctt cctgtttgttta ttacgtatg cacttctca cagtgttctg gaaggttctg	2460
tttgccttcg tcccacccatc agaataactgg aatggctggg cctgcttcat tgcctccatc	2520
ctcatgatcg gcctactgac cgcccttcatt ggagacctgg ctccccactt tggctgcacc	2580
attggctctga aagattccgt gactgccgtt gtgtttgttgc ctcttggAAC ctccggccca	2640
gacacatttg ccagcaaagt agcagctacc caggaccagt atgcagatgc gtctataggc	2700
aatgtcaactg gaagcaatgc tgcataatgtc ttccctggaa tcggcgtggc ctggccatt	2760
gctgccatct gccatgcggc caacggggaa cagttcaaaag tgcgtccccgg cacgttagct	2820
tgcctctgtca ctctcttcac tattttgttgc ttcatcaacg tgggggtgtct gctgtatcg	2880
cggaggccag aaataggagc catcttcggg atgtcaataa ttctgttagt aatttcacag	2940
agaacttccc aagaacgtct tgctgttgtt cagttcaag ttcaagcttt ttgggtgttt	3000
tgcctttca gtcttcatgc tgcgttgttgc atctgtatcg gatcttcagg gttcttaatt	3060
attatgttgc aatgatttca ttgttgttgc ttgttgttgc ttatgttgc ttatgttgc	3120
acggcttttag ggggacagggc acacttcggg ctgttcctga caattggatc acattattga	3180
cacagaaaaac aaaaacaaaaa atcaataaag ttctactg tttgtgttgc catcaccagt	3240
ggcacacgcat gttctgtact gtggaaatgtc cctcttcctca ctgtcaccta cttcatttc	3300
agttgttgttgc ttcatgatc tgaagaaata aagacaaaaa aggc	3344

<210> SEQ ID NO 16

<211> LENGTH: 940

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 16

Met Leu Arg Leu Ser Leu Pro Pro Asn Val Ser Met Gly Phe Arg Leu			
1	5	10	15

Val Ala Leu Val Ala Leu Leu Phe Ser His Val Asp His Ile Thr Ala		
20	25	30

Asp Thr Glu Ala Glu Thr Gly Gly Asn Glu Thr Thr Glu Cys Thr Gly		
35	40	45

Ser Tyr Tyr Cys Lys Lys Gly Val Ile Leu Pro Ile Trp Glu Pro Gln		
50	55	60

Asp Pro Ser Phe Gly Asp Lys Ile Ala Arg Ala Thr Val Tyr Phe Val			
65	70	75	80

Ala Met Val Tyr Met Phe Leu Gly Val Ser Ile Ile Ala Asp Arg Phe		
85	90	95

Met Ser Ser Ile Glu Val Ile Thr Ser Gln Glu Lys Glu Ile Thr Ile		
100	105	110

Lys Lys Pro Asn Gly Glu Thr Thr Lys Thr Thr Val Arg Ile Trp Asn		
115	120	125

Glu Thr Val Ser Asn Leu Thr Leu Met Ala Leu Gly Ser Ser Ala Pro		
130	135	140

Glu Ile Leu Leu Ser Val Ile Glu Val Cys Gly His Asn Phe Thr Ala			
145	150	155	160

-continued

Gly Asp Leu Gly Pro Ser Arg Ile Val Gly Ser Ala Ala Phe Asn Met
165 170 175

Phe Ile Ile Ile Ala Leu Cys Val Tyr Val Val Pro Asp Gly Glu Thr
180 185 190

Arg Lys Ile Lys His Leu Arg Val Phe Phe Val Thr Ala Ala Trp Ser
195 200 205

Ile Phe Ala Tyr Thr Trp Leu Tyr Ile Ile Leu Ser Val Ser Ser Pro
210 215 220

Gly Val Val Glu Val Trp Glu Gly Leu Leu Thr Phe Phe Phe Pro
225 230 235 240

Ile Cys Val Val Phe Ala Trp Val Ala Asp Arg Arg Leu Leu Phe Tyr
245 250 255

Lys Tyr Val Tyr Lys Arg Tyr Arg Ala Gly Lys Gln Arg Gly Met Ile
260 265 270

Ile Glu His Glu Gly Asp Arg Pro Ala Ser Lys Thr Glu Ile Glu Met
275 280 285

Asp Gly Lys Val Val Asn Ser His Val Asp Asn Phe Leu Asp Gly Ala
290 295 300

Leu Val Leu Glu Val Asp Glu Arg Asp Gln Asp Asp Glu Glu Ala Arg
305 310 315 320

Arg Glu Met Ala Arg Ile Leu Lys Glu Leu Lys Gln Lys His Pro Glu
325 330 335

Lys Glu Ile Glu Gln Leu Ile Glu Leu Ala Asn Tyr Gln Val Leu Ser
340 345 350

Gln Gln Gln Lys Ser Arg Ala Phe Tyr Arg Ile Gln Ala Thr Arg Leu
355 360 365

Met Thr Gly Ala Gly Asn Ile Leu Lys Arg His Ala Ala Asp Gln Ala
370 375 380

Arg Lys Ala Val Ser Met His Glu Val Asn Met Glu Met Ala Glu Asn
385 390 395 400

Asp Pro Val Ser Lys Ile Phe Phe Glu Gln Gly Thr Tyr Gln Cys Leu
405 410 415

Glu Asn Cys Gly Thr Val Ala Leu Thr Ile Met Arg Arg Gly Gly Asp
420 425 430

Leu Ser Thr Thr Val Phe Val Asp Phe Arg Thr Glu Asp Gly Thr Ala
435 440 445

Asn Ala Gly Ser Asp Tyr Glu Phe Thr Glu Gly Thr Val Ile Phe Lys
450 455 460

Pro Gly Glu Thr Gln Lys Glu Ile Arg Val Gly Ile Ile Asp Asp Asp
465 470 475 480

Ile Phe Glu Glu Asp Glu Asn Phe Leu Val His Leu Ser Asn Val Arg
485 490 495

Val Ser Ser Asp Val Ser Glu Asp Gly Ile Leu Glu Ser Asn His Ala
500 505 510

Ser Ser Ile Ala Cys Leu Gly Ser Pro Ser Thr Ala Thr Ile Thr Ile
515 520 525

Phe Asp Asn Asp His Thr Gly Ile Phe Thr Phe Glu Glu Pro Val Thr
530 535 540

His Val Ser Glu Ser Ile Gly Ile Met Glu Val Lys Val Leu Arg Thr
545 550 555 560

Ser Gly Ala Arg Gly Asn Val Ile Ile Pro Tyr Lys Thr Ile Glu Gly

-continued

565	570	575
Thr Ala Arg Gly Gly Glu Asp Phe Glu Asp Thr Cys Gly Val Pro		
580	585	590
Glu Phe Gln Asn Asp Glu Ile Val Lys Thr Ile Ser Val Lys Val Ile		
595	600	605
Ile Asp Asp Glu Glu Tyr Glu Lys Asn Lys Thr Phe Ile Glu Ile Gly		
610	615	620
Glu Pro Arg Leu Val Glu Met Ser Glu Lys Lys Ala Leu Leu Leu Asn		
625	630	635
640		
Glu Leu Gly Gly Phe Thr Leu Thr Gly Lys Glu Met Tyr Gly Gln Pro		
645	650	655
Ile Phe Arg Lys Val His Ala Arg Asp His Pro Ile Pro Ser Thr Val		
660	665	670
Ile Thr Ile Ser Glu Glu Tyr Asp Asp Lys Gln Pro Leu Thr Ser Lys		
675	680	685
Glu Asp Gln Glu Arg Arg Ile Ala Glu Met Gly Arg Pro Ile Leu Gly		
690	695	700
Glu His Thr Lys Leu Glu Val Ile Ile Glu Glu Ser Tyr Glu Phe Lys		
705	710	715
720		
Ser Thr Val Asp Lys Leu Ile Lys Lys Thr Asn Leu Ala Leu Val Val		
725	730	735
Gly Thr Asn Ser Trp Arg Glu Gln Phe Ile Glu Ala Ile Thr Val Ser		
740	745	750
Ala Gly Glu Arg Asp Asp Asp Glu Tyr Gly Glu Glu Lys Leu Pro		
755	760	765
Ser Cys Phe Asp Tyr Val Met His Phe Leu Thr Val Phe Trp Lys Val		
770	775	780
Leu Phe Ala Phe Val Pro Pro Thr Glu Tyr Trp Asn Gly Trp Ala Cys		
785	790	795
800		
Phe Ile Val Ser Ile Leu Met Ile Gly Leu Leu Thr Ala Phe Ile Gly		
805	810	815
Asp Leu Ala Ser His Phe Gly Cys Thr Ile Gly Leu Lys Asp Ser Val		
820	825	830
Thr Ala Val Val Phe Val Ala Leu Gly Thr Ser Val Pro Asp Thr Phe		
835	840	845
Ala Ser Lys Val Ala Ala Thr Gln Asp Gln Tyr Ala Asp Ala Ser Ile		
850	855	860
Gly Asn Val Thr Gly Ser Asn Ala Val Asn Val Phe Leu Gly Ile Gly		
865	870	875
880		
Val Ala Trp Ser Ile Ala Ala Ile Cys His Ala Ala Asn Gly Glu Gln		
885	890	895
Phe Lys Val Ser Pro Gly Thr Leu Ala Phe Ser Val Thr Leu Phe Thr		
900	905	910
Ile Phe Ala Phe Ile Asn Val Gly Val Leu Leu Tyr Arg Arg Arg Pro		
915	920	925
Glu Ile Gly Ala Ile Phe Gly Met Ser Ile Ile Leu		
930	935	940

<210> SEQ ID NO 17

<211> LENGTH: 3262

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 17

gcggcggggc tggcagcagt ggctgcccgc actgcgcccc ggccgctcgcc ttgcgtcag	60
ctcccggtgc cgccgctcg ggccggccccc cggcaggcccc tcctcggtat ggccgcggcc	120
tcctcccccgc ccagggccga gaggaagcgc tgggggttggg gccgcctgccc aggcgcggcc	180
cggggcagcg cgggcctggc caagaagtgc cccttctcg tggagctggc ggagggcggc	240
cggcgccggcg gcgcgcgtcta cgcccccata gcgcggccggc ccccgaggctc cgcccccct	300
gcgtccccgg ccgcgcggcg cgccgcggca gttgcctccg accttgcccgc gggccggcc	360
gtgagccttag accccgcgt ctccatctac agcacgcgc gcccggttggt ggccgcacc	420
cacgtccagg gccgcgtcta caacttcctc gagcgtccca cccggtggaa atgcttcgtt	480
taccacttcg ccgtcttcct catcgcttgc gtctgcctca tcttcagcgt gctgtccacc	540
atcgagcagt atgcgcgcct ggccacgggg actctttctt ggtatggagat cgtgtgggt	600
gtgttcttcg ggacggagta cgtggtcggc ctctggtcgg cccggtggcg cagcaagtac	660
gtgggcctct gggggcggct ggcgtttggc cggaaagccca ttccatcat cgacctcate	720
gtggtcgtgg cctccatggt ggttccttcg gtgggtccca agggggcagggt gtttgcacg	780
tcggccatca ggggcattccg cttcctgcag atcctgagga tgctacacgt cgaccggcag	840
ggaggcacct ggaggcgtct gggctccgtg gtcttcatcc accgcgcaggaa gctgataacc	900
accctgtaca tcggcttcct gggcctcata ttcttcctcgt actttgtgtt cctggctgag	960
aaggacgcgg tgaacgcgtc agggcgcgtg gagttcggca gctacgcaga tgcgtgtgg	1020
tgggggttgg tcacagtca caccatcggt tatggggaca aggtgcggca gacgtgggtc	1080
ggaaagacca tcgccttcgt ctcttcgttgc ttggccatct ctttcttcg gctcccaagcg	1140
gggattcttg gtcgggggtt tgccctgaag gtgcagcaga agcagaggca gaagcacttc	1200
aacccgcaga tcccgccgc agoctcactc attcagaccc catggagggt ctatgtgcc	1260
gagaaccccg acttcctccac ctggaaagatc tacatccggaa aggccccccg gagccacact	1320
ctgtgtcac ccaggcccaa acccaagaag tctgtgggtg taaagaaaaaa aaagtcaag	1380
ctggacaaag acaatgggt gactcctgga gagaagatgc tcacagtccc ccataatcag	1440
tgcgacccccc cagaagagcg gcccgtggc cacttcctcg tcgacggcta tgacagttct	1500
gtaaggaaga gccaacact gctggaaagtg agcatgcggc atttcatgag aaccaacagc	1560
ttcgccgagg acctggaccc ggaagggggactctgtca cacccatcac ccacatctca	1620
cagctgcggg aacaccatcg ggccacccatt aaggtcatc gacgcgtca gtacttgt	1680
gccaagaaga aattccagca agcgccggaa ctttacgtat tgccggacgt cattgagcag	1740
tactcgcagg gccacctcaa cttcatggtg cgcaccaagg agctgcagag gaggtggac	1800
cagtccattg ggaagccctc actgttcatc tccgtctcg aaaagagcaa ggatgcggc	1860
agcaacacga tcggcgcccg ctttacgtat tgccggacgt cattgagcag	1920
aggctggcact tcatcacca gacgtgtct ctttgcacgg tggcagcacc	1980
ccggcagcg gcccggccccc cagagaggc gggggccaca tcacccagcc ctgcggcagt	2040
ggccggctccg tcgaccctga gcttttcctcg cccagcaaca ccctgcccac ctacgagcag	2100
ctgaccctgtgc ccaggaggggg ccccgatgag gggtcctgag gagggatgg ggctggggg	2160
tgggcctgag tgagaggggaa ggccaagagt ggccccacact ggccctctct gaaggaggcc	2220

-continued

acctcctaaa	aggcccagag	agaagagccc	cactctcaga	ggccccaata	ccccatggac	2280
catgctgtct	ggcacagcct	gcacttgggg	gctcagcaag	gccaccttt	cctggccggt	2340
gtggggccc	cgtctcaggt	ctgagttgtt	accccaagcg	ccctggcccc	cacatggtga	2400
tgttgacatc	actggcatgg	tggttgggac	ccagtgccag	ggcacaggc	ctggccatg	2460
tatggccagg	aagttagcaca	ggotgagtgc	aggcccaccc	tgcttggccc	agggggcttc	2520
ctgaggggag	acagagcaac	ccctggaccc	cagcctcaaa	tccaggaccc	tgccaggcac	2580
aggcagggca	ggaccagccc	acgctgacta	cagggccgccc	ggcaataaaaa	gcccaggagc	2640
ccatttggag	ggectggcc	tggtccctc	actctcagga	aatgctgacc	catggcagg	2700
agactgtgga	gactgctct	gagccccag	cttccagcag	gagggacagt	ctcaccattt	2760
ccccagggca	cgtgggttag	tggggggaaac	gcccacttcc	ctgggtaga	ctgccagctc	2820
tccctagctg	gagaggagcc	ctgcctctcc	gcccctgagc	ccactgtgcg	tggggctccc	2880
gcctccaaacc	cctcgcccag	tcccagcgc	cagccaaaca	cacagaagg	gactgcacc	2940
tccccttgc	agctgtgag	ccgcagagaa	gtgacggtcc	ctacacagga	caggggttcc	3000
ttctggcat	tacatcgcat	agaaatcaat	aatttgggt	gatttggatc	tgtgttttaa	3060
tgagttcac	agtgtgattt	tgattattaa	tttgcaagc	tttcctaata	aaacgtggag	3120
aatcacaggc	tgggctgggc	actgctctca	cttgggttcc	tggggcatcc	atggggctcc	3180
tcacagacag	gaccctgca	gttcccctgg	aagcagtgcc	caggtggctg	tggaatagga	3240
acgctaaaaaa	aaaaaaaaaa	aa				3262

<210> SEQ ID NO 18

<211> LENGTH: 676

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

Met	Ala	Ala	Ser	Ser	Pro	Pro	Arg	Ala	Glu	Arg	Lys	Arg	Trp	Gly
1								10			15			

Trp	Gly	Arg	Leu	Pro	Gly	Ala	Arg	Arg	Gly	Ser	Ala	Gly	Leu	Ala	Lys
								20		25		30			

Lys	Cys	Pro	Phe	Ser	Leu	Glu	Leu	Ala	Glu	Gly	Gly	Pro	Ala	Gly	Gly
								35		40		45			

Ala	Leu	Tyr	Ala	Pro	Ile	Ala	Pro	Gly	Ala	Pro	Gly	Pro	Ala	Pro	Pro
								50		55		60			

Ala	Ser	Pro	Ala	Ala	Pro	Ala	Ala	Pro	Pro	Val	Ala	Ser	Asp	Leu	Gly
								65		70		75		80	

Pro	Arg	Pro	Pro	Val	Ser	Leu	Asp	Pro	Arg	Val	Ser	Ile	Tyr	Ser	Thr
								85		90		95			

Arg	Arg	Pro	Val	Leu	Ala	Arg	Thr	His	Val	Gln	Gly	Arg	Val	Tyr	Asn
								100		105		110			

Phe	Leu	Glu	Arg	Pro	Thr	Gly	Trp	Lys	Cys	Phe	Val	Tyr	His	Phe	Ala
								115		120		125			

Val	Phe	Leu	Ile	Val	Leu	Val	Cys	Leu	Ile	Phe	Ser	Val	Leu	Ser	Thr
								130		135		140			

Ile	Glu	Gln	Tyr	Ala	Ala	Leu	Ala	Thr	Gly	Thr	Leu	Phe	Trp	Met	Glu
								145		150		155		160	

Ile	Val	Leu	Val	Val	Phe	Phe	Gly	Thr	Glu	Tyr	Val	Val	Arg	Leu	Trp
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

-continued

165	170	175
Ser Ala Gly Cys Arg Ser Lys Tyr Val Gly Leu Trp Gly Arg Leu Arg		
180	185	190
Phe Ala Arg Lys Pro Ile Ser Ile Ile Asp Leu Ile Val Val Val Ala		
195	200	205
Ser Met Val Val Leu Cys Val Gly Ser Lys Gly Gln Val Phe Ala Thr		
210	215	220
Ser Ala Ile Arg Gly Ile Arg Phe Leu Gln Ile Leu Arg Met Leu His		
225	230	235
240		
Val Asp Arg Gln Gly Gly Thr Trp Arg Leu Leu Gly Ser Val Val Phe		
245	250	255
Ile His Arg Gln Glu Leu Ile Thr Thr Leu Tyr Ile Gly Phe Leu Gly		
260	265	270
Leu Ile Phe Ser Ser Tyr Phe Val Tyr Leu Ala Glu Lys Asp Ala Val		
275	280	285
Asn Glu Ser Gly Arg Val Glu Phe Gly Ser Tyr Ala Asp Ala Leu Trp		
290	295	300
Trp Gly Val Val Thr Val Thr Ile Gly Tyr Gly Asp Lys Val Pro		
305	310	315
320		
Gln Thr Trp Val Gly Lys Thr Ile Ala Ser Cys Phe Ser Val Phe Ala		
325	330	335
Ile Ser Phe Phe Ala Leu Pro Ala Gly Ile Leu Gly Ser Gly Phe Ala		
340	345	350
Leu Lys Val Gln Gln Lys Gln Arg Gln Lys His Phe Asn Arg Gln Ile		
355	360	365
Pro Ala Ala Ala Ser Leu Ile Gln Thr Ala Trp Arg Cys Tyr Ala Ala		
370	375	380
Glu Asn Pro Asp Ser Ser Thr Trp Lys Ile Tyr Ile Arg Lys Ala Pro		
385	390	395
400		
Arg Ser His Thr Leu Leu Ser Pro Ser Pro Lys Pro Lys Lys Ser Val		
405	410	415
Val Val Lys Lys Lys Phe Lys Leu Asp Lys Asp Asn Gly Val Thr		
420	425	430
Pro Gly Glu Lys Met Leu Thr Val Pro His Ile Thr Cys Asp Pro Pro		
435	440	445
Glu Glu Arg Arg Leu Asp His Phe Ser Val Asp Gly Tyr Asp Ser Ser		
450	455	460
Val Arg Lys Ser Pro Thr Leu Leu Glu Val Ser Met Pro His Phe Met		
465	470	475
480		
Arg Thr Asn Ser Phe Ala Glu Asp Leu Asp Leu Glu Gly Glu Thr Leu		
485	490	495
Leu Thr Pro Ile Thr His Ile Ser Gln Leu Arg Glu His His Arg Ala		
500	505	510
Thr Ile Lys Val Ile Arg Arg Met Gln Tyr Phe Val Ala Lys Lys Lys		
515	520	525
Phe Gln Gln Ala Arg Lys Pro Tyr Asp Val Arg Asp Val Ile Glu Gln		
530	535	540
Tyr Ser Gln Gly His Leu Asn Leu Met Val Arg Ile Lys Glu Leu Gln		
545	550	555
560		
Arg Arg Leu Asp Gln Ser Ile Gly Lys Pro Ser Leu Phe Ile Ser Val		
565	570	575

-continued

Ser	Glu	Lys	Ser	Lys	Asp	Arg	Gly	Ser	Asn	Thr	Ile	Gly	Ala	Arg	Leu
580															590
Asn	Arg	Val	Glu	Asp	Lys	Val	Thr	Gln	Leu	Asp	Gln	Arg	Leu	Ala	Leu
595															605
Ile	Thr	Asp	Met	Leu	His	Gln	Leu	Leu	Ser	Leu	His	Gly	Gly	Ser	Thr
610															620
Pro	Gly	Ser	Gly	Gly	Pro	Pro	Arg	Glu	Gly	Gly	Ala	His	Ile	Thr	Gln
625															640
Pro	Cys	Gly	Ser	Gly	Gly	Ser	Val	Asp	Pro	Glu	Leu	Phe	Leu	Pro	Ser
645															655
Asn	Thr	Leu	Pro	Thr	Tyr	Glu	Gln	Leu	Thr	Val	Pro	Arg	Arg	Gly	Pro
660															670
Asp	Glu	Gly	Ser												
															675

<210> SEQ ID NO 19
<211> LENGTH: 3052
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 19

cccggtggctg	ggctggcagc	agtggctacc	cgcgcgtgcgc	cctgcgtct	gctgcgtcctt	60
cacctcagct	ccgaggagca	gcggggtcag	gggtcctgtc	tggccatgga	cacggcctcg	120
tccccggcca	gtgctgaaag	gaagcgccgc	ggttggagcc	gcctgttagg	cgcccccgg	180
ggcagcgcgg	tggtcaagaa	gtgtcccttc	tcaactggagc	tggccgaagg	tggccctgag	240
ggcagcacgg	tctatgcgcc	catcgccca	accggagccc	ccgggctcgc	gccccccattg	300
tccgaccccg	tgtcgcccc	cccgccccct	gcagacccctcg	gcccacgtcc	gccccgtgagc	360
cttgacccgc	gggtctccat	ctacagtgc	cgccgcccc	tgctggcgc	caccacatc	420
caggggccgag	tctacaactt	cctcgagcgc	cccacgggtt	ggaagtgttt	cgtgtaccac	480
ttcaccgtct	tcctcattgt	tctggtctgc	ctcatcttca	gtgtcctgtc	cactatttag	540
cagtatgccc	ctctggccac	cgggaccctc	ttctggatgg	agattgtcct	tgtgggttcc	600
tttgggacag	aatatgtggt	ccgcctctgg	tctgcaggct	gccgcagcaa	gtacgtggc	660
atctggggcc	ggctacgttt	tgcccgaaag	cccatttcca	tcattgacct	catcgtggtt	720
gtagcctcta	tggttgtct	ctgcgtgggt	tccaaaggac	aagtgttcgc	cacatcagct	780
atcaggggta	tccgcttct	tcagatcttg	cgatgctgc	atgtcgatcg	ccaggggggt	840
acctggaggc	tcctgggctc	tgtagtcttc	attcaccgc	aggagctgat	caccacccctg	900
tacattggct	ttctgggct	tatcttctcc	tcctactttg	tctacttgc	tgagaaaagat	960
gcggtaacg	agtccggccg	catcgagttt	ggcagctacg	cagatgtct	gtgggtgggg	1020
gtggtcacag	tcactaccat	tggctacggg	gataaggtac	ctcagacgtg	ggttgggaag	1080
accatcgct	cctgtttctc	tgtcttcgcc	atatccttct	ttgcactccc	agcggggata	1140
cttggctctg	ggttcgcgt	gaaggtccag	cagaagcaga	ggcagaagca	cttcaaccgg	1200
cagatcccag	ctgcagccctc	actcatccag	actgcatgga	ggtgctatgc	cgctgagaac	1260
cctgactcag	ccacttggaa	gatctatgtc	cggaagccctg	ctcggagtc	cacgcttctg	1320
tcccccagcc	ccaaacctaa	aaagtctgtc	atggtaaaga	agaagaagtt	caagctggat	1380

-continued

aaggataatg	ggatgagttcc	tggagagaag	atgttcaatg	ttcctcacat	cacttatgtat	1440
ccccccagagg	ataggaggcc	agaccatttc	tccattgtat	gctatgcacag	ctcagtaagg	1500
aagagcccta	cactgctgga	agtaagcaca	ccccatttct	tgagaacaaa	cagcttgca	1560
gaggacctgg	accttggaaagg	ggagacactg	ctgaccccca	tcacccatgt	gtcacagctg	1620
cgggatcacc	atcgggccac	catcaaggtc	atcagggcga	tgcagtactt	tgttagccaa	1680
aagaattcc	agcaagcacf	gaagccctac	gacgtgegag	atgtcatcg	gcagttactcc	1740
cagggccacc	tgaaccttat	ggtgcgcat	aaagaactac	agagaaggct	ggatcagtcc	1800
atggaaagc	catctttgtt	catccccatc	tcaaaaaaaa	gcaaagaccg	tggcagtaac	1860
accatcggtg	cccgtctgaa	cagggtggaa	gacaagggtg	cacaactgga	ccagagactg	1920
gtgatcatca	cagacatgt	ccaccagctg	ctgtccatgc	aacaagggtt	tccaaacctgc	1980
aacagcaggt	cacaagtgt	agccagcaat	gaagggtggct	ccatcaacccc	tgagcttcc	2040
ctacccagca	acagcctgcc	cacctaegaa	caactgactg	tgcccccac	aggccctgat	2100
gagggttcct	gaaggagctc	agaagggagg	tccagggtaa	accccccaca	ggtctgcccc	2160
accccccct	caacaggggc	acacccac	ggttctcacc	cccttgaagg	cctgatatga	2220
cagectagct	ccctaaggcc	ccgaaaccca	tgggccaagc	cactggcctg	gatctggata	2280
tcaccaacac	ccctctgccc	agccagtggg	aactagaact	tgtctggca	tgggttcct	2340
ctcaggcccc	tatcaagatg	ttgatgtgc	tcttctgtat	gctggAACAT	agggatgggg	2400
agtatatgt	taaaggcagga	tgcagctcg	gcagagggaa	gcccattccag	gtatgccaag	2460
tgggttccta	aaggaaaca	ggaagagtct	ggggtttcca	ccctcagcccc	tgaatccaga	2520
gaccatcagg	gaacgtgcag	ggcaagacag	gccccggccca	cactaactgc	aggctgtac	2580
ccagtggag	ctcagcaggc	tagagcctg	ctcccttcatt	ctctagtgtat	gttgaccaca	2640
acaagatact	agcgggattt	tcacaccccc	cccaaaggat	atcccaagca	agggattgtc	2700
cactgtccac	cctagccagg	tatggactga	caacttctcc	cctaaagtga	aaagcaggcc	2760
cctcaagccc	tatgagcccc	acatgtgcag	cctgcctct	tcttgctgtt	cagaggtgt	2820
ggaccaatgc	tcagactcca	gcctcttca	cgcacattg	aagtttagggg	catcaccgat	2880
tcttccagga	ttgggttcctt	gggcttgggc	accacgcctc	acagaaatca	agataatttc	2940
tggtgattca	gactcttgt	ttttaatgaa	tttcgtatgt	tgatcttgat	tgagagactt	3000
ttccaaataa	agttggacg	tcccaggta	aaaaaaaaaa	aaaaaaaaaa	aa	3052

<210> SEQ_ID NO 20

<211> LENGTH: 668

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 20

Met	Asp	Thr	Ala	Ser	Ser	Pro	Pro	Ser	Ala	Glu	Arg	Lys	Arg	Ala	Gly
1															

5 10 15

Trp	Ser	Arg	Leu	Leu	Gly	Ala	Arg	Arg	Gly	Ser	Ala	Val	Val	Lys	Lys
20															

25 30

Cys	Pro	Phe	Ser	Leu	Glu	Leu	Ala	Glu	Gly	Gly	Pro	Glu	Gly	Ser	Thr
35															

40 45

Val	Tyr	Ala	Pro	Ile	Ala	Pro	Thr	Gly	Ala	Pro	Gly	Leu	Ala	Pro	Pro
50															

55 60

-continued

Met	Ser	Thr	Pro	Val	Ser	Pro	Ala	Pro	Ala	Asp	Leu	Gly	Pro		
65				70			75			80					
Arg	Pro	Arg	Val	Ser	Leu	Asp	Pro	Arg	Val	Ser	Ile	Tyr	Ser	Ala	Arg
	85				90				95						
Arg	Pro	Leu	Leu	Ala	Arg	Thr	His	Ile	Gln	Gly	Arg	Val	Tyr	Asn	Phe
	100			105				110							
Leu	Glu	Arg	Pro	Thr	Gly	Trp	Lys	Cys	Phe	Val	Tyr	His	Phe	Thr	Val
	115				120			125							
Phe	Leu	Ile	Val	Leu	Val	Cys	Leu	Ile	Phe	Ser	Val	Leu	Ser	Thr	Ile
	130			135			140								
Glu	Gln	Tyr	Ala	Ala	Leu	Ala	Thr	Gly	Thr	Leu	Phe	Trp	Met	Glu	Ile
	145			150			155		160						
Val	Leu	Val	Val	Phe	Phe	Gly	Thr	Glu	Tyr	Val	Val	Arg	Leu	Trp	Ser
	165			170			175								
Ala	Gly	Cys	Arg	Ser	Lys	Tyr	Val	Gly	Ile	Trp	Gly	Arg	Leu	Arg	Phe
	180			185			190								
Ala	Arg	Lys	Pro	Ile	Ser	Ile	Ile	Asp	Leu	Ile	Val	Val	Val	Ala	Ser
	195			200			205								
Met	Val	Val	Leu	Cys	Val	Gly	Ser	Lys	Gly	Gln	Val	Phe	Ala	Thr	Ser
	210			215			220								
Ala	Ile	Arg	Gly	Ile	Arg	Phe	Leu	Gln	Ile	Leu	Arg	Met	Leu	His	Val
	225			230			235		240						
Asp	Arg	Gln	Gly	Gly	Thr	Trp	Arg	Leu	Leu	Gly	Ser	Val	Val	Phe	Ile
	245			250			255								
His	Arg	Gln	Glu	Leu	Ile	Thr	Thr	Leu	Tyr	Ile	Gly	Phe	Leu	Gly	Leu
	260			265			270								
Ile	Phe	Ser	Ser	Tyr	Phe	Val	Tyr	Leu	Ala	Glu	Lys	Asp	Ala	Val	Asn
	275			280			285								
Glu	Ser	Gly	Arg	Ile	Glu	Phe	Gly	Ser	Tyr	Ala	Asp	Ala	Leu	Trp	Trp
	290			295			300								
Gly	Val	Val	Thr	Val	Thr	Thr	Ile	Gly	Tyr	Gly	Asp	Lys	Val	Pro	Gln
	305			310			315		320						
Thr	Trp	Val	Gly	Lys	Thr	Ile	Ala	Ser	Cys	Phe	Ser	Val	Phe	Ala	Ile
	325			330			335								
Ser	Phe	Phe	Ala	Leu	Pro	Ala	Gly	Ile	Leu	Gly	Ser	Gly	Phe	Ala	Leu
	340			345			350								
Lys	Val	Gln	Gln	Lys	Gln	Arg	Gln	Lys	His	Phe	Asn	Arg	Gln	Ile	Pro
	355			360			365								
Ala	Ala	Ala	Ser	Leu	Ile	Gln	Thr	Ala	Trp	Arg	Cys	Tyr	Ala	Ala	Glu
	370			375			380								
Asn	Pro	Asp	Ser	Ala	Thr	Trp	Lys	Ile	Tyr	Val	Arg	Lys	Pro	Ala	Arg
	385			390			395		400						
Ser	His	Thr	Leu	Leu	Ser	Pro	Ser	Pro	Lys	Pro	Lys	Lys	Ser	Val	Met
	405			410			415								
Val	Lys	Lys	Lys	Phe	Lys	Leu	Asp	Lys	Asp	Asn	Gly	Met	Ser	Pro	
	420			425			430								
Gly	Glu	Lys	Met	Phe	Asn	Val	Pro	His	Ile	Thr	Tyr	Asp	Pro	Pro	Glu
	435			440			445								
Asp	Arg	Arg	Pro	Asp	His	Phe	Ser	Ile	Asp	Gly	Tyr	Asp	Ser	Ser	Val
	450			455			460								
Arg	Lys	Ser	Pro	Thr	Leu	Leu	Glu	Val	Ser	Thr	Pro	His	Phe	Leu	Arg

-continued

465	470	475	480
Thr Asn Ser Phe Ala Glu Asp Leu Asp Leu Glu Gly Glu Thr Leu Leu			
485	490	495	
Thr Pro Ile Thr His Val Ser Gln Leu Arg Asp His His Arg Ala Thr			
500	505	510	
Ile Lys Val Ile Arg Arg Met Gln Tyr Phe Val Ala Lys Lys Lys Phe			
515	520	525	
Gln Gln Ala Arg Lys Pro Tyr Asp Val Arg Asp Val Ile Glu Gln Tyr			
530	535	540	
Ser Gln Gly His Leu Asn Leu Met Val Arg Ile Lys Glu Leu Gln Arg			
545	550	555	560
Arg Leu Asp Gln Ser Ile Gly Lys Pro Ser Leu Phe Ile Pro Ile Ser			
565	570	575	
Glu Lys Ser Lys Asp Arg Gly Ser Asn Thr Ile Gly Ala Arg Leu Asn			
580	585	590	
Arg Val Glu Asp Lys Val Thr Gln Leu Asp Gln Arg Leu Val Ile Ile			
595	600	605	
Thr Asp Met Leu His Gln Leu Leu Ser Met Gln Gln Gly Gly Pro Thr			
610	615	620	
Cys Asn Ser Arg Ser Gln Val Val Ala Ser Asn Glu Gly Gly Ser Ile			
625	630	635	640
Asn Pro Glu Leu Phe Leu Pro Ser Asn Ser Leu Pro Thr Tyr Glu Gln			
645	650	655	
Leu Thr Val Pro Gln Thr Gly Pro Asp Glu Gly Ser			
660	665		

<210> SEQ ID NO 21
<211> LENGTH: 3900
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

ccatgggctc	aggatgccgg	tgcggagggg	ccacgtcgcg	ccgcagaaca	ccttccttgg	60
caccatcatc	cgcaagtttg	agggccagag	ccgtaagttc	atcatcgcca	acgctcggtt	120
ggagaactgc	gccgtcatct	actgcaacga	cggcttctgc	gagctgtgcg	gctactcgcg	180
ggccgaggtg	atgcagcgcac	cctgcacctg	cgacttcttg	cacggggccgc	gcacgcagcg	240
ccgcgtgtcc	gcccgcgcac	cgccaggactg	gctgggcgcc	gaggagcgca	aagtggaaat	300
cgccttctac	cggaaagatg	ggagctgttt	cctatgtctg	gtggatgtgg	tgccctgtaa	360
gaacgaggat	ggggctgtca	tcatgttcat	cctcaatttc	gaggtggtga	tggagaagga	420
catgggtgggg	tccccggctc	atgacaccaa	ccaccggggc	ccccccacca	gtgggtggc	480
cccaaggccgc	gccaagaccc	tccgcctgaa	gtgccccgcg	ctgctggcgc	tgaaggcccc	540
ggagtcgtcg	gtgcgggtcg	ggggcgccgg	gggcgcgggc	gccccggggg	ccgtgggtgg	600
ggacgtggac	ctgacgcggc	cggcacccag	cagcgagtcg	ctggccctgg	acgaagtgc	660
agccatggac	aaccacgtgg	cagggctcg	gcccgcggag	gagcgccgtg	cgctgggtgg	720
tcccggtctc	ccgcggccca	gcccgcgggg	ccagctccca	tccgcggccgg	cgcacagcct	780
caaccccgac	gcctcggtct	ccagctcgag	cctggccccc	acgcgcctcc	gagaaagctg	840
cgccagcggt	cgcccgccct	cgtcgccgca	cgacatcgag	gccatgcgcg	ccgggggtgt	900

-continued

ggccccgcca	ccgcgccacg	ccagcacccgg	ggccatgcac	ccactgcgca	gcgggttgct	960
caactccacc	tcggactccg	acctcgtgca	ctaccgcacc	attagcaaga	ttccccaaat	1020
cacccctcaac	tttgtggacc	tcaagggcga	ccccttcttg	gcttcgccc	ccagtgaccg	1080
tgagatcata	gcacctaaga	taaaggagcg	aacccacaat	gtcactgaga	aggtcaccca	1140
ggtccctgtcc	ctgggcgcgg	acgtgctgcc	ttagtacaag	ctgcaggcac	cgcgcatcca	1200
ccgtggacc	atcctgcatt	acagccccct	caaggccgtg	tgggactggc	tcatectgct	1260
gctggtcata	tacacggctg	tcttcacacc	ctactcggt	gccttcctgc	tgaaggagac	1320
ggaagaaggc	ccgctctgta	ccgagtggtgg	ctacgcctgc	cagccgcgtgg	ctgtggtgga	1380
cctcatcg	gacatcatgt	tcattgtgga	catcctcatc	aacttccgca	ccacctacgt	1440
caatgccaac	gaggaggtgg	ttagccaccc	cgccgcac	gcccgtccact	acttcaaggg	1500
ctgggttcctc	atcgacatgg	tggccgcac	ccccttcgac	ctgctcatct	tcggctctgg	1560
ctctgaggag	ctgatcgggc	tgctgaagac	tgcgcggctg	ctgcggctgg	tgcgcgtggc	1620
gcggaaagctg	gatcgctact	cagagtacgg	cgccgcgtg	ctgttctgc	tcatgtgcac	1680
ctttgcgtc	atcgcgact	ggctagectg	catctggta	gccatcgca	acatggagca	1740
gccacacatg	gactcacgca	tcggctggct	gcacaacctg	ggcgaccaga	taggcaaacc	1800
ctacaacagc	agcggccctgg	ggggccccc	catcaaggac	aagtatgtga	cgccgcgtcta	1860
cttcacccctc	agcagccctca	ccagtggtgg	cttcggcaac	gtctctccca	acaccaactc	1920
agagaagatc	ttctccatct	gctgtatgt	cattggctcc	ctcatgtatg	ctagcatctt	1980
cggaacacgt	tcggccatca	tccagcggt	gtactcggtc	acagccccct	accacacaca	2040
gatgctgcgg	gtgcgggagt	tcatccgctt	ccaccagatc	cccaatcccc	tgcgcagcg	2100
cctcgaggag	tacttccagc	acgcctggtc	ctacaccaac	ggcatcgaca	tgaacgcgg	2160
gctgaagggc	ttccctgagt	gcctgcaggc	tgacatctgc	ctgcacctga	accgctcact	2220
gctgcagcac	tgcaaaccct	tccgaggggc	caccaagggc	tgccttcggg	ccctggccat	2280
gaagttcaag	accacacatg	caccgcagg	ggacacactg	gtgcgtctg	gggacctgct	2340
caccgcctg	tacttcatct	cccgcccc	catcgagatc	ctgcggggc	acgtcgctg	2400
ggccatctg	gggaagaatg	acatctttgg	ggagcctctg	aacctgtatg	caaggcttgg	2460
caagtgcac	ggggatgtgc	ggggccctac	ctactgtgac	ctacacaa	tccatcgga	2520
cgacctgctg	gaggtgctgg	acatgtaccc	ttagttctcc	gaccacttct	ggtccagcct	2580
ggagatcacc	ttcaacctgc	gagataccaa	catgatcccg	ggctcccccg	gcagtacgga	2640
gttagagggt	ggtttcagtc	ggcaacgcac	gcaagttg	tccctccgca	ggcgacacgg	2700
caaggacacg	gacgagccag	gggaggtgtc	ggcttgggg	ccggggccggg	cgggggcagg	2760
gccgagtagc	cgggcccccgc	cgggggggcc	gtggggggag	agcccggtca	gtggccccc	2820
cagccctgag	agcagtgagg	atgagggccc	aggccgcac	tccagcccc	tccgcctgtt	2880
gcccttcac	agccccagc	ccccggaga	gcccggggt	ggggagcccc	tgatggagga	2940
ctgcgagaag	agcagcgaca	cttgcaccc	cctgtcaggc	gccttcctcg	gagtgtccaa	3000
catttcagc	ttctgggggg	acagtcgggg	ccgcccgtac	caggagctcc	ctcgatgccc	3060
cgccccccacc	cccaagccctcc	tcaacatccc	cctctccagc	ccgggtcgcc	ggccccgggg	3120
cgacgtggag	agcaggctgg	atgcctccca	gcccgcac	aacaggctgg	agaccggcgt	3180

-continued

gagtcagac atggccactg tctgcagct gctacagagg cagatgacgc tggteccgcc	3240
cgcctacagt gctgtgacca ccccggggcc tggccccact tccacatccc cgctgtgcc	3300
cgtcagcccc ctccccaccc tcacccitgga ctgcgtttct caggtttccc agtcatggc	3360
gtgtgaggag ctgccccegg gggccccaga gttcccaa gaagggccca cacgacgcct	3420
ctccctaccc ggccagctgg gggccctcac ctcccgccctc acggctcgga	3480
ccggggcagt tagtggggct gcccagtgtg gacacgtggc tcaccaggg atcaaggcgc	3540
tgctggggcg ctccccctgg agggccctgct caggaggccc tgaccgtgga aggggagagg	3600
aactcgaaag cacagctct ccccccagcccttgggaccat cttctctgc agtccccctgg	3660
gccccagtgaa gaaaaaaa gggggcagg ggcagggccg gcagtaggtg gggcctgtgg tccccccact	3720
gccccctggg cattagctgg tctaactgcc cggaggcacc cggccctggg ccttaggcac	3780
ctcaaggact ttctgttat ttactgtct tattgttaag gataataatt aaggatcata	3840
tgaataatta atgaagatgc tgatgactat gaataataaa taattatcct gaggagaaaa	3900

<210> SEQ ID NO 22

<211> LENGTH: 1159

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22

Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Phe Leu Asp			
1	5	10	15

Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Ile Ile Ala		
20	25	30

Asn Ala Arg Val Glu Asn Cys Ala Val Ile Tyr Cys Asn Asp Gly Phe		
35	40	45

Cys Glu Leu Cys Gly Tyr Ser Arg Ala Glu Val Met Gln Arg Pro Cys		
50	55	60

Thr Cys Asp Phe Leu His Gly Pro Arg Thr Gln Arg Arg Ala Ala Ala			
65	70	75	80

Gln Ile Ala Gln Ala Leu Leu Gly Ala Glu Glu Arg Lys Val Glu Ile		
85	90	95

Ala Phe Tyr Arg Lys Asp Gly Ser Cys Phe Leu Cys Leu Val Asp Val		
100	105	110

Val Pro Val Lys Asn Glu Asp Gly Ala Val Ile Met Phe Ile Leu Asn		
115	120	125

Phe Glu Val Val Met Glu Lys Asp Met Val Gly Ser Pro Ala His Asp		
130	135	140

Thr Asn His Arg Gly Pro Pro Thr Ser Trp Leu Ala Pro Gly Arg Ala			
145	150	155	160

Lys Thr Phe Arg Leu Lys Leu Pro Ala Leu Leu Ala Leu Thr Ala Arg		
165	170	175

Glu Ser Ser Val Arg Ser Gly Gly Ala Gly Gly Ala Gly Ala Pro Gly		
180	185	190

Ala Val Val Val Asp Val Asp Leu Thr Pro Ala Ala Pro Ser Ser Glu		
195	200	205

Ser Leu Ala Leu Asp Glu Val Thr Ala Met Asp Asn His Val Ala Gly		
210	215	220

Leu Gly Pro Ala Glu Glu Arg Arg Ala Leu Val Gly Pro Gly Ser Pro			
225	230	235	240

-continued

Pro Arg Ser Ala Pro Gly Gln Leu Pro Ser Pro Arg Ala His Ser Leu
 245 250 255
 Asn Pro Asp Ala Ser Gly Ser Ser Cys Ser Leu Ala Arg Thr Arg Ser
 260 265 270
 Arg Glu Ser Cys Ala Ser Val Arg Arg Ala Ser Ser Ala Asp Asp Ile
 275 280 285
 Glu Ala Met Arg Ala Gly Val Leu Pro Pro Pro Arg His Ala Ser
 290 295 300
 Thr Gly Ala Met His Pro Leu Arg Ser Gly Leu Leu Asn Ser Thr Ser
 305 310 315 320
 Asp Ser Asp Leu Val Arg Tyr Arg Thr Ile Ser Lys Ile Pro Gln Ile
 325 330 335
 Thr Leu Asn Phe Val Asp Leu Lys Gly Asp Pro Phe Leu Ala Ser Pro
 340 345 350
 Thr Ser Asp Arg Glu Ile Ile Ala Pro Lys Ile Lys Glu Arg Thr His
 355 360 365
 Asn Val Thr Glu Lys Val Thr Gln Val Leu Ser Leu Gly Ala Asp Val
 370 375 380
 Leu Pro Glu Tyr Lys Leu Gln Ala Pro Arg Ile His Arg Trp Thr Ile
 385 390 395 400
 Leu His Tyr Ser Pro Phe Lys Ala Val Trp Asp Trp Leu Ile Leu Leu
 405 410 415
 Leu Val Ile Tyr Thr Ala Val Phe Thr Pro Tyr Ser Ala Ala Phe Leu
 420 425 430
 Leu Lys Glu Thr Glu Glu Gly Pro Pro Ala Thr Glu Cys Gly Tyr Ala
 435 440 445
 Cys Gln Pro Leu Ala Val Val Asp Leu Ile Val Asp Ile Met Phe Ile
 450 455 460
 Val Asp Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Ala Asn Glu
 465 470 475 480
 Glu Val Val Ser His Pro Gly Arg Ile Ala Val His Tyr Phe Lys Gly
 485 490 495
 Trp Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu Ile
 500 505 510
 Phe Gly Ser Gly Ser Glu Glu Leu Ile Gly Leu Leu Lys Thr Ala Arg
 515 520 525
 Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr Ser Glu
 530 535 540
 Tyr Gly Ala Ala Val Leu Phe Leu Leu Met Cys Thr Phe Ala Leu Ile
 545 550 555 560
 Ala His Trp Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Met Glu Gln
 565 570 575
 Pro His Met Asp Ser Arg Ile Gly Trp Leu His Asn Leu Gly Asp Gln
 580 585 590
 Ile Gly Lys Pro Tyr Asn Ser Ser Gly Leu Gly Gly Pro Ser Ile Lys
 595 600 605
 Asp Lys Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser
 610 615 620
 Val Gly Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Ile Phe
 625 630 635 640

-continued

Ser Ile Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe
645 650 655

Gly Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg
660 665 670

Tyr His Thr Gln Met Leu Arg Val Arg Glu Phe Ile Arg Phe His Gln
675 680 685

Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His Ala
690 695 700

Trp Ser Tyr Thr Asn Gly Ile Asp Met Asn Ala Val Leu Lys Gly Phe
705 710 715 720

Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn Arg Ser Leu
725 730 735

Leu Gln His Cys Lys Pro Phe Arg Gly Ala Thr Lys Gly Cys Leu Arg
740 745 750

Ala Leu Ala Met Lys Phe Lys Thr Thr His Ala Pro Pro Gly Asp Thr
755 760 765

Leu Val His Ala Gly Asp Leu Leu Thr Ala Leu Tyr Phe Ile Ser Arg
770 775 780

Gly Ser Ile Glu Ile Leu Arg Gly Asp Val Val Val Ala Ile Leu Gly
785 790 795 800

Lys Asn Asp Ile Phe Gly Glu Pro Leu Asn Leu Tyr Ala Arg Pro Gly
805 810 815

Lys Ser Asn Gly Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu His Lys
820 825 830

Ile His Arg Asp Asp Leu Leu Glu Val Leu Asp Met Tyr Pro Glu Phe
835 840 845

Ser Asp His Phe Trp Ser Ser Leu Glu Ile Thr Phe Asn Leu Arg Asp
850 855 860

Thr Asn Met Ile Pro Gly Ser Pro Gly Ser Thr Glu Leu Glu Gly Gly
865 870 875 880

Phe Ser Arg Gln Arg Lys Arg Lys Leu Ser Phe Arg Arg Arg Thr Asp
885 890 895

Lys Asp Thr Glu Gln Pro Gly Glu Val Ser Ala Leu Gly Pro Gly Arg
900 905 910

Ala Gly Ala Gly Pro Ser Ser Arg Gly Arg Pro Gly Gly Pro Trp Gly
915 920 925

Glu Ser Pro Ser Ser Gly Pro Ser Ser Pro Glu Ser Ser Glu Asp Glu
930 935 940

Gly Pro Gly Arg Ser Ser Ser Pro Leu Arg Leu Val Pro Phe Ser Ser
945 950 955 960

Pro Arg Pro Pro Gly Glu Pro Pro Gly Gly Glu Pro Leu Met Glu Asp
965 970 975

Cys Glu Lys Ser Ser Asp Thr Cys Asn Pro Leu Ser Gly Ala Phe Ser
980 985 990

Gly Val Ser Asn Ile Phe Ser Phe Trp Gly Asp Ser Arg Gly Arg Gln
995 1000 1005

Tyr Gln Glu Leu Pro Arg Cys Pro Ala Pro Thr Pro Ser Leu Leu Asn
1010 1015 1020

Ile Pro Leu Ser Ser Pro Gly Arg Arg Pro Arg Gly Asp Val Glu Ser
1025 1030 1035 1040

Arg Leu Asp Ala Leu Gln Arg Gln Leu Asn Arg Leu Glu Thr Arg Leu

-continued

1045	1050	1055	
Ser Ala Asp Met Ala Thr Val Leu Gln Leu Leu Gln Arg Gln Met Thr			
1060	1065	1070	
Leu Val Pro Pro Ala Tyr Ser Ala Val Thr Thr Pro Gly Pro Gly Pro			
1075	1080	1085	
Thr Ser Thr Ser Pro Leu Leu Pro Val Ser Pro Leu Pro Thr Leu Thr			
1090	1095	1100	
Leu Asp Ser Leu Ser Gln Val Ser Gln Phe Met Ala Cys Glu Glu Leu			
1105	1110	1115	1120
Pro Pro Gly Ala Pro Glu Leu Pro Gln Glu Gly Pro Thr Arg Arg Leu			
1125	1130	1135	
Ser Leu Pro Gly Gln Leu Gly Ala Leu Thr Ser Gln Pro Leu His Arg			
1140	1145	1150	
His Gly Ser Asp Pro Gly Ser			
1155			

<210> SEQ ID NO 23

<211> LENGTH: 4235

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 23

gaattcccg	gagccggcg	ctggagcg	ggagcgccg	cgcggggacg	agggtgcagcc	60
gggcccgtgg	caccaccaggc	ccgggtcgac	agacaggtgt	gccggcgccg	cagctccagc	120
agtttcaggc	ccggccagg	gtccggagcc	ggcagccact	aagcagacg	ggaatccccgc	180
gccctcggtc	cagttccggc	cccgccccca	ctcagcttgg	ccgcggggtg	cgagaccac	240
ggcccgccca	ggccacccgca	agoctagtgc	tggggccggc	cgggccaggg	tgggtggggg	300
cccgcccgcc	ccgccccatgg	gctcaggatg	ccgggtcg	ggggccacgt	cgcgcgcgcag	360
aacaccttcc	tcgacacccat	catccgcaag	tttggggcc	agagccgca	gttcatcatc	420
gctaacgcgc	gcgttagagaa	ctgcgctgtc	atctactgca	acgacggctt	ctgcgaactg	480
tgtggtaact	cggggccgca	ggtgtatgc	cgccctgtca	cctgcgtt	cctgcgtatgg	540
ccgcgcacgc	agegcccgtgc	cgccgcgcag	atcgcgagg	ccctgttgg	cgcagaggag	600
cgcaaagtgg	agatgcctt	ctaccggaaa	gatgggagct	gcttcctgt	tctgggtggat	660
gtggtaacct	tgaagaatga	agatggggct	gtgtatcatgt	tcatcctcaa	ctttgaagta	720
gtgtatggaga	aggacatgg	agggtccccc	gctcatgaca	ccaaccacag	ggggccctct	780
accagctggc	tagttctgg	ccgggccaag	actttccccc	tgaagctgcc	tgccttgg	840
gcactgacgg	ccagggaaatc	atctgtcg	acaggaagca	tgcgcgtgc	tggagccct	900
ggggccgggtgg	tgggtggatgt	ggacctgacg	ccggcagcac	ccagcagtga	gtccctggcc	960
ttggatgagg	tttctgccc	ggacaaccat	gtggcagg	ttgggtgtgc	agaagagagg	1020
cgagcactgg	tggggccggg	gtctgctca	ccagtagcca	gcatccgagg	ccctcaccca	1080
tcgcccacgag	ctcagacccat	taaccctgtat	gcctcagg	ccagctgc	cctggcccg	1140
acacgctccc	gtgagagctg	cgctagegtg	cgccgcgc	catccgcca	tgcatttg	1200
gcaatgcggg	ctggagcgct	gccccctccg	ccccgcctat	caagcacagg	ggccatgcac	1260
ccctgcgc	gtggcctgtct	taactccacc	tcaagactctg	accttgc	ctaccgaacc	1320
attagcaaga	tacccaaat	cactctcaac	tttgcgacc	tcaaaggcga	ccctttcctg	1380

-continued

gcttcaccca ccagtgaccg ggagattata gcacccaaga taaaagagcg gaccacaat	1440
gtcaccgaga aggtcaccca ggtcctgtcc ttgggagcag atgtgctgcc tgagtataag	1500
ctgcaggccc caagaatcca ccgctggacc atcctccact acagccccctt caaggcggtg	1560
tgggactggc tcatcctgct gctggcgtac tacacggcag tcttcacacc gtactcggcc	1620
gccttcctgc tgaaggagac tgaagatggg tcccaagccc ctgactgtgg ctacgctgc	1680
cagccccctga ctgttagtggc cctcatcgta gacatcatgt tcattgtggc catccttac	1740
aattccgtta ccacctatgt taatgccaac gaggaggtgg tttagccatcc tggccgcatt	1800
gctgtgcact acttcaaggg ctgggttctc attgacatgg tggctgccc catcttcgac	1860
ctgctcatct ttggctctgg ctccgaggag ctgatcgggc tgctgaaaaac tgcgaggctg	1920
ctgcggctgg tgccgcgtgc tcgaaagctg gaccgctact cggagttacgg ggcagcagtg	1980
ctttccctgc tcatgtgcac ctgcgcctc attgcgcact gggtggctg catctggta	2040
gcgcgcgca acatggaca gcttcacatg gactcacaca tcggctggct gcacaacttg	2100
ggcgaccaga tccggcaagcc ctacaadage agcggcctgg gaggcccctc catcaaggac	2160
aagtatgtca cagcgctcta cttcacccctc agcagcctca ccagcgtggg ctttggcaat	2220
gtctctccca acaccaactc agagaagatc ttctccatct gtgtcatgt cattggctcc	2280
ctcatgtatg ccgcgcattt cgccaaacgtg tccgcacatca tccagcggtt gtactcggc	2340
acggcccgct accacacaca gatgctccgg gtgcggggagt ttatccgctt tcaccagatt	2400
cccaacccat tacggcagcg cctcgaggag tacttccagc atgcttggtc ctacaccaat	2460
ggcatcgaca tgaacgcggc gctgaagggc ttccctgagt gctccaggc tgacatctgc	2520
ctacacctga acgcgcgtct gctgcgcgtat tgcaagccat tccgaggggc cactaaaggc	2580
tacctgaggg cgctggctat gaagttcaag accacacatg caccaccagg ggacacacta	2640
gtgcacgcgg gggacctgct tactgcgcctc tacttcatct ccaggggctc catcgagatc	2700
ctgcgggggtg atgtcgtggt ggccatctt gggaaagaacg acatcttgg agagectctg	2760
aaccttatatg ccegtcctgg aaagtccaaat ggggatgtgc gggccctcact gtaactgtac	2820
ctgcacaaga tccatcgta tgacttgcgtt gaggtgttgg acatgtatcc cgagttctca	2880
gatcaactttt ggtctagect ggagatcacc ttcaacccctc gagataccaa catgattcct	2940
ggctccccag gcagtgcaga actagagagc ggcttaaca ggcaacgcaa ggcacaaactg	3000
tcattccgca ggcgtacaga caaggacaca gagcagccag gggagggtgc agccctgggg	3060
caaggccctg cccgagttgg gccaggggcc agttgcggg gacagccggg agggccatgg	3120
ggggagagcc catccagegg cccttccagc ccagagagca gtgaagatga gggccctgg	3180
cgcagctcca gtcggctccg cctgggtgcc ttctccagcc ccaggectcc gggagacccc	3240
ccaggtgggg agectttgac agaggatggg gagaaaatg acacctgtaa tcccttgtca	3300
ggtgcccttc ctgggggtgtc caatatttc agttctggg ggaacagtcg gggggccag	3360
taccaggagt tgccctcgatc ccctgcggcc tccccccagcc tcctcaacat ccccttgtct	3420
agccctggc ggcgcattccg gggtgatgtg gagagcagcc tggacgcact ccagagacag	3480
ctgaacaggg tggaaaccccg gctaagtgcac gacatggcca ctgtcctaca gctgctacag	3540
aggcaaatga ccctggtccc tcctgcctac agtgcgttaa ccacccctgg gccccggcccc	3600
acttccgcattt cccctttgtt gctgttcggc cctgtccccca ctctcaccct ggactcgctt	3660

-continued

tctcagggtt cccagttcggt ggcttttagag gagcttcctg caggagcccc agagctcccc	3720
caagatggcc ccactcgacg cctctccctg cccggccagc tggggccct tacctcccag	3780
ccccctgcaca gacatggctc agatccaggc agttagtgtat gctgcccact gtggacatgt	3840
ggctcacccct gggttcaggg cattacctga ccgctctcct tggagaccct ggtttgagg	3900
ggagacctgt gcagaagcac ctccatcccc acccttggga ccagcatctc ctcctgcaat	3960
accctgggcc ccactgggag gggcaggaca gagctggca gtggatgggg gtctgtggtc	4020
cccccaactgc cctgggggca ttagctggtc taactgcccc gaggcacccgg ccctggccct	4080
taggcacctc aaggactttt ctgctattta ctgcttttat tgtaaggat aataattaag	4140
gatcatatga ataattaatg aagatgctga tgactatgaa taataaataa ttatcctgag	4200
qaqactcccaq qqtqctqqttt aqcttcaaaa aaaaa	4235

<210> SEQ ID NO 24

<211> LENGTH: 1162

<212> TYPE: PRT

<212> TYPE: PRI
<213> ORGANISM: *Mus musculus*

<400> SEQUENCE: 24

Met	Pro	Val	Arg	Arg	Gly	His	Val	Ala	Pro	Gln	Asn	Thr	Phe	Leu	Asp
1				5					10					15	

Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Ile Ile Ala
20 25 30

Asn Ala Arg Val Glu Asn Cys Ala Val Ile Tyr Cys Asn Asp Gly Phe
 35 40 45

Cys Glu Leu Cys Gly Tyr Ser Arg Ala Glu Val Met Gln Arg Pro Cys
 50 55 60

Gln Ile Ala Gln Ala Leu Leu Gly Ala Glu Glu Arg Lys Val Glu Ile
85 90 95

Ala Phe Tyr Arg Lys Asp Gly Ser Cys Phe Leu Cys Leu Val Asp Val
 100 105 110

Val Pro Val Lys Asn Glu Asp Gly Ala Val Ile Met Phe Ile Leu Asn
115 120 125

Phe Glu Val Val Met Glu Lys Asp Met Val Gly Ser Pro Ala His Asp
130 135 140

Thr Asn His Arg Gly Pro Ser Thr Ser Trp Leu Ala Ser Gly Arg Ala
145 150 155 160

Lys Thr Phe Arg Leu Lys Leu Pro Ala Leu Leu Ala Leu Thr Ala Arg
165 170 175

Glu Ser Ser Val Arg Thr Gly Ser Met Arg Ser Ala Gly Ala Pro Gly
180 185 190

Ala Val Val Val Asp Val Asp Leu Thr Pro Ala Ala Pro Ser Ser Glu
195 200 205

Ser Leu Ala Leu Asp Glu Val Ser Ala Met Asp Asn His Val Ala Gly
210 215 220

Leu Gly Pro Ala Glu Glu Arg Arg Ala Leu Val Gly Pro Gly Ser Ala
225 230 235 240

Ser Pro Val Ala Ser Ile Arg Gly Pro His Pro Ser Pro Arg Ala Gln
245 250 255

-continued

Ser Leu Asn Pro Asp Ala Ser Gly Ser Ser Cys Ser Leu Ala Arg Thr
 260 265 270
 Arg Ser Arg Glu Ser Cys Ala Ser Val Arg Arg Ala Ser Ser Ala Asp
 275 280 285
 Asp Ile Glu Ala Met Arg Ala Gly Ala Leu Pro Pro Pro Pro Arg His
 290 295 300
 Ala Ser Thr Gly Ala Met His Pro Leu Arg Ser Gly Leu Leu Asn Ser
 305 310 315 320
 Thr Ser Asp Ser Asp Leu Val Arg Tyr Arg Thr Ile Ser Lys Ile Pro
 325 330 335
 Gln Ile Thr Leu Asn Phe Val Asp Leu Lys Gly Asp Pro Phe Leu Ala
 340 345 350
 Ser Pro Thr Ser Asp Arg Glu Ile Ile Ala Pro Lys Ile Lys Glu Arg
 355 360 365
 Thr His Asn Val Thr Glu Lys Val Thr Gln Val Leu Ser Leu Gly Ala
 370 375 380
 Asp Val Leu Pro Glu Tyr Lys Leu Gln Ala Pro Arg Ile His Arg Trp
 385 390 395 400
 Thr Ile Leu His Tyr Ser Pro Phe Lys Ala Val Trp Asp Trp Leu Ile
 405 410 415
 Leu Leu Leu Val Ile Tyr Thr Ala Val Phe Thr Pro Tyr Ser Ala Ala
 420 425 430
 Phe Leu Leu Lys Glu Thr Glu Asp Gly Ser Gln Ala Pro Asp Cys Gly
 435 440 445
 Tyr Ala Cys Gln Pro Leu Thr Val Val Asp Leu Ile Val Asp Ile Met
 450 455 460
 Phe Ile Val Asp Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Ala
 465 470 475 480
 Asn Glu Glu Val Val Ser His Pro Gly Arg Ile Ala Val His Tyr Phe
 485 490 495
 Lys Gly Trp Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu
 500 505 510
 Leu Ile Phe Gly Ser Gly Ser Glu Glu Leu Ile Gly Leu Leu Lys Thr
 515 520 525
 Ala Arg Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr
 530 535 540
 Ser Glu Tyr Gly Ala Ala Val Leu Phe Leu Leu Met Cys Thr Phe Ala
 545 550 555 560
 Leu Ile Ala His Trp Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Met
 565 570 575
 Glu Gln Pro His Met Asp Ser His Ile Gly Trp Leu His Asn Leu Gly
 580 585 590
 Asp Gln Ile Gly Lys Pro Tyr Asn Ser Ser Gly Leu Gly Pro Ser
 595 600 605
 Ile Lys Asp Lys Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu
 610 615 620
 Thr Ser Val Gly Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu Lys
 625 630 635 640
 Ile Phe Ser Ile Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser
 645 650 655

-continued

Ile Phe Gly Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr
660 665 670

Ala Arg Tyr His Thr Gln Met Leu Arg Val Arg Glu Phe Ile Arg Phe
675 680 685

His Gln Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln
690 695 700

His Ala Trp Ser Tyr Thr Asn Gly Ile Asp Met Asn Ala Val Leu Lys
705 710 715 720

Gly Phe Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn Arg
725 730 735

Ser Leu Leu Gln His Cys Lys Pro Phe Arg Gly Ala Thr Lys Gly Tyr
740 745 750

Leu Arg Ala Leu Ala Met Lys Phe Lys Thr Thr His Ala Pro Pro Gly
755 760 765

Asp Thr Leu Val His Ala Gly Asp Leu Leu Thr Ala Leu Tyr Phe Ile
770 775 780

Ser Arg Gly Ser Ile Glu Ile Leu Arg Gly Asp Val Val Val Ala Ile
785 790 795 800

Leu Gly Lys Asn Asp Ile Phe Gly Glu Pro Leu Asn Leu Tyr Ala Arg
805 810 815

Pro Gly Lys Ser Asn Gly Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu
820 825 830

His Lys Ile His Arg Asp Asp Leu Leu Glu Val Leu Asp Met Tyr Pro
835 840 845

Glu Phe Ser Asp His Phe Trp Ser Ser Leu Glu Ile Thr Phe Asn Leu
850 855 860

Arg Asp Thr Asn Met Ile Pro Gly Ser Pro Gly Ser Ala Glu Leu Glu
865 870 875 880

Ser Gly Phe Asn Arg Gln Arg Lys Arg Lys Leu Ser Phe Arg Arg Arg
885 890 895

Thr Asp Lys Asp Thr Glu Gln Pro Gly Glu Val Ser Ala Leu Gly Gln
900 905 910

Gly Pro Ala Arg Val Gly Pro Gly Pro Ser Cys Arg Gly Gln Pro Gly
915 920 925

Gly Pro Trp Gly Glu Ser Pro Ser Ser Gly Pro Ser Ser Pro Glu Ser
930 935 940

Ser Glu Asp Glu Gly Pro Gly Arg Ser Ser Ser Pro Leu Arg Leu Val
945 950 955 960

Pro Phe Ser Ser Pro Arg Pro Pro Gly Asp Pro Pro Gly Gly Glu Pro
965 970 975

Leu Thr Glu Asp Gly Glu Lys Ser Asp Thr Cys Asn Pro Leu Ser Gly
980 985 990

Ala Phe Ser Gly Val Ser Asn Ile Phe Ser Phe Trp Gly Asn Ser Arg
995 1000 1005

Gly Arg Gln Tyr Gln Glu Leu Pro Arg Cys Pro Ala Pro Ala Pro Ser
1010 1015 1020

Leu Leu Asn Ile Pro Leu Ser Ser Pro Gly Arg Arg Ser Arg Gly Asp
1025 1030 1035 1040

Val Glu Ser Arg Leu Asp Ala Leu Gln Arg Gln Leu Asn Arg Leu Glu
1045 1050 1055

Thr Arg Leu Ser Ala Asp Met Ala Thr Val Leu Gln Leu Gln Arg

-continued

1060	1065	1070	
Gln Met Thr Leu Val Pro Pro Ala Tyr Ser Ala Val Thr Thr Pro Gly			
1075	1080	1085	
Pro Gly Pro Thr Ser Ala Ser Pro Leu Leu Pro Val Gly Pro Val Pro			
1090	1095	1100	
Thr Leu Thr Leu Asp Ser Leu Ser Gln Val Ser Gln Phe Val Ala Phe			
1105	1110	1115	1120
Glu Glu Leu Pro Ala Gly Ala Pro Glu Leu Pro Gln Asp Gly Pro Thr			
1125	1130	1135	
Arg Arg Leu Ser Leu Pro Gly Gln Leu Gly Ala Leu Thr Ser Gln Pro			
1140	1145	1150	
Leu His Arg His Gly Ser Asp Pro Gly Ser			
1155	1160		

<210> SEQ ID NO 25

<211> LENGTH: 1901

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25

gggcaggaag acggcgctgc ccggaggagc ggggcgggcg ggccgcgggg ggagcggg	60
gcggcgccgg gccaggcccc ggcggggggcg gggggggggc ggcagaaga ggccggggc	120
cgcgtccgg cccgtctgcg cggttggcct tggctttggc ttggcgccg cggttgaga	180
agatgctgca gtccctggcc ggcagctcg gcgtgcgcct ggtggagcg caccgtcg	240
cctggtgctt cggcttcgt gtgctggcct acttgctcta cctggcttc ggccactgg	300
tcttcctcctc ggtggagctg ccctatgggg acctgctgcg ccaggagctg cgcaagctg	360
agcgacgcctt ctggaggag cacgagtgcc tgtctgagca gcagctggag cagttccctg	420
gcccgggtgct ggaggccagc aactacggcg tgcgggtgct cagcaacgcc tcgggcaact	480
ggaactggga cttcacctcc gcgcctttct tcgcccagcac cgtgcctcc accacagg	540
atggccacac cgtgccttg tcagatggag gtaaggcctt ctgcatacatc tactccgtca	600
ttggcattcc cttcacccctc ctgttccatc cggctgtggt ccagcgcatac accgtgcac	660
tcacccgcag gccggctctc tacttccaca tccgctgggg cttctccaag caggtggtag	720
ccatcgtccca tgccgtgtc cttgggtttc tcactgtgtc ctgcattttc ttcatccgg	780
ccgctgtctt ctcagtcctg gaggatgact ggaacttcctt ggaatccccc tattttgtt	840
ttatccctt gggcaccatt ggcctggggg attatgtgcc tggggaaaggc tacaatcaa	900
aattcagaga gctctataag attgggatca cgtgttacct gctacttggc cttattgcc	960
tgttgttagt tctggaaacc ttctgtgaac tccatgagct gaaaaaatcc agaaaaatgt	1020
tctatgtcaa gaaggacaag gacgaggatc aggtgcacat catagagcat gaccaactgt	1080
ccttcctcctc gatcacagac caggcagctg gcatgaaaga ggaccagaag caaaatgag	1140
ctttgtggc cacccagtca tctgcctgcg tggatggccc tgcaaaccat tgagcgtagg	1200
atttgttgcg ttagtgcata gcaccagggt cagggtgcaa ggaagaggct taagtatgtt	1260
cattttatc agaatgcaaa agcgaaaatt atgtcactt aagaaatagc tactgtttgc	1320
aatgtcttat taaaaaaacaa caaaaaaaga cacatggAAC aaagaagctg tgacccagc	1380
aggatgtcta atatgtgagg aaatgagatg tccacctaattcataatgt gacaaaatta	1440

-continued

tctcgaccc tt acataggagg agaataacttg aagcagtatg ctgcgtgggt tagaagcaga	1500
ttttatactt ttaactggaa actttgggg ttgcatttag atcatttagc tgatggctaa	1560
atagcaaaaat ttatatttag aagcaaaaaa aaaaagcata gagatgtgtt ttataaataag	1620
gtttatgtgt actgggttgc atgtaccac caaaaatgtat tattttggaa gaatctaagt	1680
caaactcaactt atttataatg cataggtaac catthaactat gtacatataa agtataaata	1740
tgtttataatt ctgtacataat ggtttaggtc accagatcct agtgtagttc tgaaactaag	1800
actatagata ttttgggttct tttgatttct ctttataacta aagaatccag agttgtaca	1860
ataaaaataag gggataataa aacttgagag tgaataacca t	1901

<210> SEQ ID NO 26

<211> LENGTH: 336

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26

Met Leu Gln Ser Leu Ala Gly Ser Ser Cys Val Arg Leu Val Glu Arg			
1	5	10	15

His Arg Ser Ala Trp Cys Phe Gly Phe Leu Val Leu Gly Tyr Leu Leu			
20	25	30	

Tyr Leu Val Phe Gly Ala Val Val Phe Ser Ser Val Glu Leu Pro Tyr			
35	40	45	

Glu Asp Leu Leu Arg Gln Glu Leu Arg Lys Leu Lys Arg Arg Phe Leu			
50	55	60	

Glu Glu His Glu Cys Leu Ser Glu Gln Gln Leu Glu Gln Phe Leu Gly			
65	70	75	80

Arg Val Leu Glu Ala Ser Asn Tyr Gly Val Ser Val Leu Ser Asn Ala			
85	90	95	

Ser Gly Asn Trp Asn Trp Asp Phe Thr Ser Ala Leu Phe Phe Ala Ser			
100	105	110	

Thr Val Leu Ser Thr Thr Gly Tyr His Thr Val Pro Leu Ser Asp			
115	120	125	

Gly Gly Lys Ala Phe Cys Ile Ile Tyr Ser Val Ile Gly Ile Pro Phe			
130	135	140	

Thr Leu Leu Phe Leu Thr Ala Val Val Gln Arg Ile Thr Val His Val			
145	150	155	160

Thr Arg Arg Pro Val Leu Tyr Phe His Ile Arg Trp Gly Phe Ser Lys			
165	170	175	

Gln Val Val Ala Ile Val His Ala Val Leu Leu Gly Phe Val Thr Val			
180	185	190	

Ser Cys Phe Phe Ile Pro Ala Ala Val Phe Ser Val Leu Glu Asp			
195	200	205	

Asp Trp Asn Phe Leu Glu Ser Phe Tyr Phe Cys Phe Ile Ser Leu Ser			
210	215	220	

Thr Ile Gly Leu Gly Asp Tyr Val Pro Gly Glu Gly Tyr Asn Gln Lys			
225	230	235	240

Phe Arg Glu Leu Tyr Lys Ile Gly Ile Thr Cys Tyr Leu Leu Gly			
245	250	255	

Leu Ile Ala Met Leu Val Val Leu Glu Thr Phe Cys Glu Leu His Glu			
260	265	270	

Leu Lys Lys Phe Arg Lys Met Phe Tyr Val Lys Lys Asp Lys Asp Glu	
---	--

-continued

275	280	285
-----	-----	-----

Asp Gln Val His Ile Ile Glu His Asp Gln Leu Ser Phe Ser Ser Ile	290	295
		300

Thr Asp Gln Ala Ala Gly Met Lys Glu Asp Gln Lys Gln Asn Glu Pro	305	310
		315
		320

Phe Val Ala Thr Gln Ser Ser Ala Cys Val Asp Gly Pro Ala Asn His	325	330
		335

<210> SEQ ID NO 27

<211> LENGTH: 1589

<212> TYPE: DNA

<213> ORGANISM: Xenopus tropicalis

<400> SEQUENCE: 27

caaggagtga ggcagccat aacgactttg agggcaaagg gcagacatac accttgctga	60
acacaccaggc aactcccacg gcaaacacaa gtggtcctt ctctgtgggg cggccagtcg	120
ttgcttggc tgacctttgc tacctgagaa gggaggaaac cgacgtcaca atgtacagga	180
gacactgcca catcagctct ccgttcccag agagacgcta tggccccctc gcagagccgt	240
gaattttcca taacataatc aggaaatgat gctccaagcc ctggccacca acccctgcct	300
gcccgtgatc caacggAACC agacatgtcg gtgctttgtc ctcttgatgt tgggtacact	360
gctttccctt ctgatagggg cagccatatt ctcaegggtg gagctgcccc acgaaacacgt	420
tctccgggag gagctgctgg acttaaagca tcgctacctg caggagaatg agtgcctgag	480
cgaggagagg ctggagagct tcctcagccg agtactggag gccagcaatt acgggggtgc	540
catgctgaac aatgtctctg ggaaccccaa ctgggacttc acctctgccc tggctttgt	600
tagcacggtg ctctccacta cagggtatgg gcacacagtg ccgcctctcta acgcaggaa	660
gaccttctgt atcatttact ctatcatogg gatccctctg acactgtcc tggctactgc	720
gctgggtgcag cggatcatgg tgcacgtgac ccacaggccc atttcctact tccaccccg	780
ctgggggtac aacaaggcaga ctgtggccgt cgttcacgctg ctggtcatttggc	840
cattttgtgc ttttcctca taccggccgc aattttctcg gctctegaag atgattggaa	900
tttccctagag tcgttttact tctgcttcat ttctttgagc accatcgcc tgggagacta	960
tgtgcccgcg gagggggcaaa accagaggta cggccagctg tacaagtctg gcatcaact	1020
ttacctgata ttgggcctca ttgtgatgct tgggttctg gagactttct gtgagctgca	1080
ggggctgaag aaattccgaa aaatgttcta caggaagaaa atgaaagagg gagaccaatt	1140
aaacatcata gaacacgatc agctgaccct cgcctccatt tcagagcaag cggcgtcctt	1200
aaaggaggaa caaatgctgg acgaacactc tggcccgat ggcccaactaa atagtgatta	1260
gttaggacaga ctggcgctcg ttatggatcc catataattat atatactgtc tcattttca	1320
ttaaaggaga gggtctgatt aaaaagaact tatggcgtga caaggacggg ggtgttccga	1380
tcccatgtgc agttgacctg cattttattt ttattgaact ttttatttcc tagtgatgtc	1440
tttggtttag tgggttcttag cgtaacttcc cctagcaaca tacctaaggaaatgaatga	1500
ataggaacta gagcagagaa atgaaagaga aagtgattag caataaaagc ctcaccacta	1560
aaaaaaaaaaaa aaaaaaaaaa aaaaaaagg	1589

<210> SEQ ID NO 28

<211> LENGTH: 330

-continued

<212> TYPE: PRT

<213> ORGANISM: Xenopus tropicalis

<400> SEQUENCE: 28

Met Leu Gln Ala Leu Ala Thr Asn Pro Cys Leu Arg Leu Ile Gln Arg
1 5 10 15

Asn Gln Thr Cys Trp Cys Phe Val Leu Leu Met Leu Gly Tyr Leu Leu
20 25 30

Phe Leu Leu Ile Gly Ala Ala Ile Phe Ser Ala Val Glu Leu Pro His
35 40 45

Glu His Val Leu Arg Glu Glu Leu Leu Asp Leu Lys His Arg Tyr Leu
50 55 60

Gln Glu Asn Glu Cys Leu Ser Glu Glu Arg Leu Glu Ser Phe Leu Ser
65 70 75 80

Arg Val Leu Glu Ala Ser Asn Tyr Gly Val Ser Met Leu Asn Asn Val
85 90 95

Ser Gly Asn Pro Asn Trp Asp Phe Thr Ser Ala Leu Phe Phe Val Ser
100 105 110

Thr Val Leu Ser Thr Thr Gly Tyr His Thr Val Pro Leu Ser Asn
115 120 125

Ala Gly Lys Thr Phe Cys Ile Ile Tyr Ser Ile Ile Gly Ile Pro Leu
130 135 140

Thr Leu Leu Leu Phe Thr Ala Leu Val Gln Arg Ile Met Val His Val
145 150 155 160

Thr His Arg Pro Ile Ser Tyr Phe His Leu Arg Trp Gly Tyr Asn Lys
165 170 175

Gln Thr Val Ala Val Val His Ala Leu Val Ile Gly Phe Val Ala Ile
180 185 190

Leu Cys Phe Phe Leu Ile Pro Ala Ala Ile Phe Ser Ala Leu Glu Asp
195 200 205

Asp Trp Asn Phe Leu Glu Ser Phe Tyr Phe Cys Phe Ile Ser Leu Ser
210 215 220

Thr Ile Gly Leu Gly Asp Tyr Val Pro Ala Glu Gly Gln Asn Gln Arg
225 230 235 240

Tyr Arg Gln Leu Tyr Lys Phe Gly Ile Thr Cys Tyr Leu Ile Leu Gly
245 250 255

Leu Ile Val Met Leu Val Val Leu Glu Thr Phe Cys Glu Leu Gln Gly
260 265 270

Leu Lys Lys Phe Arg Lys Met Phe Tyr Arg Lys Lys Met Lys Glu Gly
275 280 285

Asp Gln Leu Asn Ile Ile Glu His Asp Gln Leu Thr Leu Ala Ser Ile
290 295 300

Ser Glu Gln Ala Ala Ser Leu Lys Glu Glu Gln Met Leu Asp Glu His
305 310 315 320

Ser Val Pro Asp Gly Pro Leu Asn Ser Asp
325 330

<210> SEQ_ID NO 29

<211> LENGTH: 5397

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 29

-continued

gcgcaactgga	gcctggeca	gchgacagcc	ttccccggcgc	cggcgcccgtg	ggtcttggga	60
attctggttt	gctttggc	actcgctttt	tacaaaccac	tggatcttac	atgcctctgt	120
accccccact	tccactccat	gtccccatgc	tcctgcgc	caacacaggac	atgttctctg	180
gatgtcagct	gagtcat	aaactctgt	catgtcagta	gacagac	cttggtaga	240
caaggctccc	agagacaccc	atctctc	cttgc	tttgg	tgtgtgtgc	300
attcaaaact	gtttctccaa	agcg	tttgc	aaaaactc	actgtttcc	360
cactggagtc	cccagcagaa	gcatggc	gtgtgc	caacc	gtac	420
tttcagaaga	agacgg	tgat	tttgc	ccatgg	caat	480
ggaagagtaa	agtccacacc	cgaca	acagt	gaggag	ccgtt	540
actgtatgt	tca	gttgc	aatgt	gggtt	agaagg	600
tcaccacgtg	tgttgc	acatt	cgcttgc	ggatgt	gtgttgc	660
tcctgtcatg	gcttgc	ttttt	ggctgt	tttgc	tttgc	720
tggatgc	catc	caa	agggc	aaagctt	gttgc	780
tccttc	tgc	tttgc	tttgc	tttgc	tttgc	840
aatgccaat	tgttgc	tttgc	atgg	tttgc	tttgc	900
cttc	tgc	tttgc	atgg	tttgc	tttgc	960
ttgttgc	atgg	tttgc	tttgc	tttgc	tttgc	1020
gagtgg	caa	tttgc	tttgc	tttgc	tttgc	1080
aatcc	gaa	tttgc	tttgc	tttgc	tttgc	1140
gtttg	acag	tttgc	tttgc	tttgc	tttgc	1200
tagatg	aaa	tttgc	tttgc	tttgc	tttgc	1260
aaatc	gtt	tttgc	tttgc	tttgc	tttgc	1320
gcttttat	ct	tttgc	tttgc	tttgc	tttgc	1380
agaag	act	tttgc	tttgc	tttgc	tttgc	1440
ctcc	tttgc	tttgc	tttgc	tttgc	tttgc	1500
cattt	gct	tttgc	tttgc	tttgc	tttgc	1560
gagttcc	caga	tttgc	tttgc	tttgc	tttgc	1620
gtgtac	ctt	tttgc	tttgc	tttgc	tttgc	1680
tctctgg	aat	tttgc	tttgc	tttgc	tttgc	1740
gatgac	ggta	tttgc	tttgc	tttgc	tttgc	1800
atgg	tttca	tttgc	tttgc	tttgc	tttgc	1860
tgtgac	ccat	tttgc	tttgc	tttgc	tttgc	1920
tgtt	tttca	tttgc	tttgc	tttgc	tttgc	1980
tgttctt	tttgc	tttgc	tttgc	tttgc	tttgc	2040
aattt	tttgc	tttgc	tttgc	tttgc	tttgc	2100
ttc	tttgc	tttgc	tttgc	tttgc	tttgc	2160
aaagtgg	caa	tttgc	tttgc	tttgc	tttgc	2220
aattc	gtt	tttgc	tttgc	tttgc	tttgc	2280

-continued

aacactagec	gaatggtagc	ctctggggtt	ttgtttttt	cttttcctcc	atgatgttaa	2340
tgggttatct	caaattttaa	gtttaactac	ctaaaataaa	taccaaagat	aatgcattt	2400
tttgcacagt	ggagcttaca	ctaaaaagaa	aacaaagccc	catgggctgc	cttgaatca	2460
agagacaata	actttaacc	ttagcaagac	cttgaaccgc	cggttcattt	tgcacccat	2520
tcagaaaata	gagcatcata	ctcaccgagt	ctagtcagt	tagtgcattt	aaaaattttg	2580
tccttcatg	taactttttt	attttaagag	gaagaagaag	aaaggggcac	acacacacaa	2640
tacggacgtc	tatccttcc	tgctaggcag	tgctggccag	gctcatgtt	agtgtgcgag	2700
atggtgatgt	actcttatat	tttctggc	ttttcctttt	gcacattcca	aaattcattt	2760
cataagacaa	gatcttcata	ggacctcctt	ggcatcctgg	cattctcaaa	actgagccat	2820
ccagcatgaa	agataaatgg	gtttaaaccc	ttgctgtga	atttattgcc	tggactgtca	2880
ggacatcacc	ageccacett	caccttaggg	aagatgccac	acctggcctc	cacactgtct	2940
cttctgatca	gtctgtctgg	attgagtctt	acagtgtcag	ataggggcgc	aaatgccaa	3000
gcagggaaac	agggagggtgt	ggacaagcca	gtttgatgca	gcacttcaga	tcaagtgcct	3060
aggaaggaga	ggaaacttgc	ctttttatg	gcagaggata	gtaatgaaaa	tgtctcagta	3120
ttttagggtc	aatgagagcc	ataaaaatat	aacataatca	caagtaaagg	agataatgg	3180
ctaaacacgc	tatttcctt	ttctgtgtc	atacttatga	ctgaatgt	gctaagcatt	3240
ttctcctgtg	gagccctaga	gcaggttact	aaggaaggac	acattgttt	ccagaagcct	3300
ccccctgcctg	gctgactgcc	ttgctagaaa	cataattttt	ttttctc	tcaagtc当地	3360
taatggaact	ctttttttt	tttttttaa	tttaaagttc	cctatttgc	atttctggg	3420
ttactgactt	tttttttaa	ttggagtctc	aaaatcaact	ctcttgcgtt	atttatatctc	3480
tgtatgcctat	aaaaaaaaacag	cttgcgtttag	aatcatgtat	tttgc当地act	gatgtttgt	3540
atggtctctg	gttcttgaac	agccatatct	gaatgccgt	cctgcaaaac	tatgacaatt	3600
tttgctgttt	tcagccttca	gatttgcgtt	cttggaaac	tgagggttta	ttttcaatga	3660
aacaaagaaa	gagatgttaa	gcaagtggtt	gttttagatc	caaatgtaaa	ggcagggtt	3720
ggaagggtgtt	taaagagttt	gaggaatttgg	ggatttgagg	gtaaagaaaa	cttacagaag	3780
aggcaacaat	ttgggtcttg	acagtgc当地	gatatttgc	gcttcagctg	ctgctattat	3840
gatgttttgc	aaaggaaaat	aatcaaacc	aagagtatc	agtgtatgt	aaattaaatg	3900
aagatacagt	ggagaatggg	ggtgaccaca	aaagaggctc	cccctaaaca	cacagtgc当地	3960
ccacttaaaa	agacttgaga	aatttgc当地	gggggtggta	tggggggggc	aagaaagagg	4020
gagggaaatc	tttcaactta	tttctgaaaa	agagaaaaaa	atataaaatt	tctggc当地	4080
agggtttgtt	tttcaagaaa	attttgc当地	agctatgtt	ttaaagtgt	cattttataa	4140
agtttatcag	atattttcat	atttaaagcc	aatgtaaaat	agagggtc当地	aaagaaaaat	4200
aattgccata	gaaagtataa	tttgc当地	gtaatttgc	agagctagta	ccttatatgt	4260
accgggttagc	atgggtttag	caaataata	ccagcctt	aagggtcgta	ttgctatgtt	4320
tttctgttat	tttgc当地	atggactgtt	catttgc当地	ctttttctag	ttattagcgt	4380
tttaacagtt	acaagctta	aatggcaatt	ttttttttt	ttttttttt	ttttttttt	4440
ttgtcaagag	ccaagacaca	ggtatgcac	gacattgatt	gctgc当地	acctcaaaa	4500
tatttgtcct	tattgactgg	gttccctaa	ttaatgtaca	catgtcatta	aatgcagac	4560

-continued

ggaggggact caccatgaat atctgggtt gattcccaga tgtgtgtgc ttctctattg	4620
caagcagatt ccctgttgg a ttacttcgg atttattccc ttttaaagaa ttttgccca	4680
tatctggaa ggcactata ttttggagg agccatagat tcctggat cctattttt	4740
aacaaaatgt agacaaatgt aactctattt tgattattga gaaaggagta gtttttatac	4800
cctctaagag tatacttcaa tcagacattt taaggatgtc actatggcac tggtgtcatt	4860
tccaaattcc tagaaaagtt tgttttactt tgttttattt ctgttaatgc attcttctt	4920
ctctttactt cctttcttac cagtacactc ctatctcaac tctgtttattt tgatgagttc	4980
tgtcccgtaa atcatattt ccttacaatt aataaatgtc acttcatattt ttataataaa	5040
ccactcagta aaagcaaaag cttgtcctga gaagtagagt gagtttttt tcactctgtg	5100
tctaataatg ttaaggtggg aaaaaaaaaa gtgtggcata gctacctgcc catccccaac	5160
cctcagcaaa gtagaatctc ttttctggta attttgggtt tccgctctgg gctctggcaa	5220
gttgaacaat cctagccatt gacaatcgta atagtttattt ttttccatt tgctgtctt	5280
ttgtatctaa agtcttccta ttgtactgca caaaccatgg attgtacata tttttatata	5340
ttatgtctta ttttatttt tctaaataaaa aaaattaaaa attgaaaaaa aaaaaaa	5397

<210> SEQ ID NO 30

<211> LENGTH: 427

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 30

Met Gly Ser Val Arg Thr Asn Arg Tyr Ser Ile Val Ser Ser Glu Glu			
1	5	10	15

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

Gly Lys Ser Lys Val His Thr Arg Gln Gln Cys Arg Ser Arg Phe Val		
35	40	45

Lys Lys Asp Gly His Cys Asn Val Gln Phe Ile Asn Val Gly Glu Lys		
50	55	60

Gly Gln Arg Tyr Leu Ala Asp Ile Phe Thr Thr Cys Val Asp Ile Arg			
65	70	75	80

Trp Arg Trp Met Leu Val Ile Phe Cys Leu Ala Phe Val Leu Ser Trp		
85	90	95

Leu Phe Phe Gly Cys Val Phe Trp Leu Ile Ala Leu Leu His Gly Asp		
100	105	110

Leu Asp Ala Ser Lys Glu Gly Lys Ala Cys Val Ser Glu Val Asn Ser		
115	120	125

Phe Thr Ala Ala Phe Leu Phe Ser Ile Glu Thr Gln Thr Thr Ile Gly		
130	135	140

Tyr Gly Phe Arg Cys Val Thr Asp Glu Cys Pro Ile Ala Val Phe Met			
145	150	155	160

Val Val Phe Gln Ser Ile Val Gly Cys Ile Ile Asp Ala Phe Ile Ile		
165	170	175

Gly Ala Val Met Ala Lys Met Ala Lys Pro Lys Lys Arg Asn Glu Thr		
180	185	190

Leu Val Phe Ser His Asn Ala Val Ile Ala Met Arg Asp Gly Lys Leu		
195	200	205

Cys Leu Met Trp Arg Val Gly Asn Leu Arg Lys Ser His Leu Val Glu	
---	--

-continued

210	215	220	
Ala His Val Arg Ala Gln Leu Leu Lys Ser Arg Ile Thr Ser Glu Gly			
225	230	235	240
Glu Tyr Ile Pro Leu Asp Gln Ile Asp Ile Asn Val Gly Phe Asp Ser			
245	250	255	
Gly Ile Asp Arg Ile Phe Leu Val Ser Pro Ile Thr Ile Val His Glu			
260	265	270	
Ile Asp Glu Asp Ser Pro Leu Tyr Asp Leu Ser Lys Gln Asp Ile Asp			
275	280	285	
Asn Ala Asp Phe Glu Ile Val Val Ile Leu Glu Gly Met Val Glu Ala			
290	295	300	
Thr Ala Met Thr Thr Gln Cys Arg Ser Ser Tyr Leu Ala Asn Glu Ile			
305	310	315	320
Leu Trp Gly His Arg Tyr Glu Pro Val Leu Phe Glu Glu Lys His Tyr			
325	330	335	
Tyr Lys Val Asp Tyr Ser Arg Phe His Lys Thr Tyr Glu Val Pro Asn			
340	345	350	
Thr Pro Leu Cys Ser Ala Arg Asp Leu Ala Glu Lys Lys Tyr Ile Leu			
355	360	365	
Ser Asn Ala Asn Ser Phe Cys Tyr Glu Asn Glu Val Ala Leu Thr Ser			
370	375	380	
Lys Glu Glu Asp Asp Ser Glu Asn Gly Val Pro Glu Ser Thr Ser Thr			
385	390	395	400
Asp Thr Pro Pro Asp Ile Asp Leu His Asn Gln Ala Ser Val Pro Leu			
405	410	415	
Glu Pro Arg Pro Leu Arg Arg Glu Ser Glu Ile			
420	425		

<210> SEQ ID NO 31

<211> LENGTH: 4274

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 31

gagagcgcac	tggagccctg	gccagcgcgc	agccttccca	gcaccggcaa	gcagggtctt	60
ggaaattctc	acttgcttcg	gctcattctc	tttcacaaaa	accactggat	cttacatgct	120
tctgtaatcc	ccacttccac	tccatgtccc	catgatcctg	taccagcaac	aggacaagtt	180
ctctggatgt	cagctgagtt	actaaggtaa	ctttgcttgt	caaaagaacc	ccaaggttct	240
cggaaagcatc	catctctctt	cattaataaa	tatataatatt	aattataatat	atataataatt	300
ttttttggtg	tgtcttcacc	gaacattcaa	aactgtttct	tctaagggtt	ttgcaaaaac	360
tcagactgtt	ttctaaagca	gaaacactgg	cgtccccagc	ggaagcaatg	ggcagtgtga	420
gaaccaacccg	ctacagcatac	gtctcttcgg	aggaagatgg	catgaagctg	gccactatgg	480
cagttgccaa	tggctttggg	aatggcaaga	gtaaagtcca	tacccgacaa	cagtgcagga	540
gccgctttgt	gaagaaagat	ggtcattgca	atgttcagtt	tatcaacgtg	ggtgagaagg	600
gacagaggta	cctggcagac	atctttacta	cctgtgtcga	catccgctgg	aggtggatgc	660
tggtatctt	ctgtcttgcc	ttcgtgtct	cctggctgtt	cttggctgt	gtgttttgt	720
tgatagccct	gctccatggg	gatctagata	cttctaaagt	gagcaaagca	tgcgtgtcag	780
aggtaaacag	cttcacggct	gccttcctct	tctccatcga	gaccaggaca	accattggct	840

-continued

atggtttcag	gtgtgtgaca	gacgagtgcc	caattgcgtt	cttcatggtg	gtattccagt	900
caatcgtagg	ctgcattcatt	gacgcctca	tcattggtgc	agtcatggcg	aagatggcaa	960
agccaaagaa	gagaatgag	actcttgtct	tcaagtacaa	tgctgtgatt	gccatgaggg	1020
atggcaact	ctgcttgatg	tggagagtgg	gtaacctcg	aaagaggccac	cttggaaag	1080
ctcatgtccg	ggcacagctt	ctcaaatcta	ggatcacttc	agaaggggag	tataccctt	1140
tggaccagat	agacatcaat	gttggttttg	atagtggaaat	tgaccgata	tttctagtgt	1200
ccccatcac	tatcgttcac	gaaatagatg	aagacagccc	tttatatgac	ttgagtaagc	1260
aggacattga	caatgcagac	tttggaaattg	ttgtcataact	ggaaggcgtg	gtggaggcga	1320
ctgccatgac	aactcaatgc	cggagttcg	atctggccaa	tgaaattctc	tggggtcacc	1380
gctatgagcc	agtgcctttt	gaagagaaac	actactataa	agtagactat	tcaagattcc	1440
ataagactta	tgaagtacct	aacacccccc	tttggtagtgc	cagagactta	gcagagaaga	1500
aatacatcct	ttcaaatgca	aattcatttt	gtatgaaaa	tgaagttgcc	ctaacaagca	1560
aagaggaaga	ggaggatagt	gagaacggag	tcccagagag	cacaaggaca	gactcacctc	1620
ctggcataga	tctccacaac	caggcaagcg	tacctctaga	gcccaggccc	ttaaggcgg	1680
aatcggagat	atgactggct	gattccgtct	ttggaataact	tactttgcta	cacagctgaa	1740
cgttggtcag	agggtccgaga	cagttataaca	gaccatggta	ctggcgaga	gggtgggtgaa	1800
agcaaggcgc	cacaagagac	taaggctagc	acaaagggtt	caaggaaaga	ctaagctgga	1860
tgactgatgt	aaagtgcctt	gcaggcctcc	aagagacatg	atggcacata	tctgttttag	1920
tataagttat	ggggttttta	atgtattgtt	ttgtgttttt	acaaaacttg	aatatgcagg	1980
caaggcctcag	tttgggtaca	tgacttacat	ggaatgtttc	tcttttagggg	aacaagagt	2040
attttaatgg	cataaacacag	gcaagactct	gccttaattt	tttggaaagc	tgctaactac	2100
atgaacacga	actgtatttt	tgttgcagt	tagtttatct	tttacataac	gttaagacgt	2160
cagtgtttag	cattgttga	agcgcacagt	gtgcattaaa	gcatcaagta	tttggctatt	2220
aactgccccaa	aatgaaagcc	tgattttctg	aggccagtaa	tttggttgt	aaaaattgtat	2280
ctctctgtct	acctgtcagt	ctctgtctct	gtctctctct	cgttatctct	ctctctctct	2340
ctctctctct	ctctctcatt	acataatagc	attataaac	actagccaaa	ttgttagcctc	2400
ttgggtttat	actttctttt	tccacaatgc	taataggta	tctcaaactt	tcagtttagac	2460
gaccttaaat	gaataccaaa	gataatgcac	attgtttgt	ttttgtttt	gttttttct	2520
ttttcttttt	ttttctttcc	tttttttttc	ttttttgtca	catggagtt	atatttttt	2580
aaaaagaaac	aaaggctct	gcacagattg	tcttggaaat	ccagagacac	catctttgaa	2640
cttcagcagg	gtgtgaagca	gtccgttcat	ttttgcacta	tagtctgaaa	agaacaccaa	2700
agtctgatca	atgttagtgt	ttttcgagtc	tgttcttttg	tgcttcttc	ttatggag	2760
gggaggaggg	aagcataggt	cacagttaca	tatgcactat	aatcaatatt	ttttcttcc	2820
ccttttctg	agctttctc	tcttagcaca	ttccaaagtt	ggttccataa	gatcagatct	2880
tggttgaact	ttgttggtaa	cctggcacct	tccaaactgt	cacccagcag	ggaagatgaa	2940
cagattcaa	atttgcagc	tgactcatag	ccttgactgc	cagggcattt	ctagctcagt	3000
gtcacattaa	gaaagatgca	ttggcctctag	tcttgcctt	ctaaaggcttc	tgcagagata	3060
gagagagaga	gagagagaga	gagagagaga	gagagagaga	gagagagaga	ctttcaatgc	3120

-continued

<210> SEO ID NO 32

<211> LENGTH: 428

<212> TYPE: PRT

<212> TITLE: TRI
<213> ORGANISM: *Mus musculus*

<400> SEQUENCE: 32

Met Gly Ser Val Arg Thr Asn Arg Tyr Ser Ile Val Ser Ser Glu Glu
1 5 10 15

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

Gly Lys Ser Lys Val His Thr Arg Gln Gln Cys Arg Ser Arg Phe Val
 35 40 45

Lys Lys Asp Gly His Cys Asn Val Gln Phe Ile Asn Val Gly Glu Lys
50 55 60

Gly Gln Arg Tyr Leu Ala Asp Ile Phe Thr Thr Cys Val Asp Ile Arg
65 70 75 80

Trp Arg Trp Met Leu Val Ile Phe Cys Leu Ala Phe Val Leu Ser Trp
85 90 95

Leu Phe Phe Gly Cys Val Phe Trp Leu Ile Ala Leu Leu His Gly Asp
100 105 110

Leu Asp Thr Ser Lys Val Ser Lys Ala Cys Val Ser Glu Val Asn Ser
115 120 125

-continued

Tyr Gly Phe Arg Cys Val Thr Asp Glu Cys Pro Ile Ala Val Phe Met
145 150 155 160

Val Val Phe Gln Ser Ile Val Gly Cys Ile Ile Asp Ala Phe Ile Ile
165 170 175

Gly Ala Val Met Ala Lys Met Ala Lys Pro Lys Lys Arg Asn Glu Thr
180 185 190

Leu Val Phe Ser His Asn Ala Val Ile Ala Met Arg Asp Gly Lys Leu
195 200 205

Cys Leu Met Trp Arg Val Gly Asn Leu Arg Lys Ser His Leu Val Glu
210 215 220

Ala His Val Arg Ala Gln Leu Leu Lys Ser Arg Ile Thr Ser Glu Gly
225 230 235 240

Glu Tyr Ile Pro Leu Asp Gln Ile Asp Ile Asn Val Gly Phe Asp Ser
245 250 255

Gly Ile Asp Arg Ile Phe Leu Val Ser Pro Ile Thr Ile Val His Glu
260 265 270

Ile Asp Glu Asp Ser Pro Leu Tyr Asp Leu Ser Lys Gln Asp Ile Asp
275 280 285

Asn Ala Asp Phe Glu Ile Val Val Ile Leu Glu Gly Met Val Glu Ala
290 295 300

Thr Ala Met Thr Thr Gln Cys Arg Ser Ser Tyr Leu Ala Asn Glu Ile
305 310 315 320

Leu Trp Gly His Arg Tyr Glu Pro Val Leu Phe Glu Glu Lys His Tyr
325 330 335

Tyr Lys Val Asp Tyr Ser Arg Phe His Lys Thr Tyr Glu Val Pro Asn
340 345 350

Thr Pro Leu Cys Ser Ala Arg Asp Leu Ala Glu Lys Lys Tyr Ile Leu
355 360 365

Ser Asn Ala Asn Ser Phe Cys Tyr Glu Asn Glu Val Ala Leu Thr Ser
370 375 380

Lys Glu Glu Glu Asp Ser Glu Asn Gly Val Pro Glu Ser Thr Ser
385 390 395 400

Thr Asp Ser Pro Pro Gly Ile Asp Leu His Asn Gln Ala Ser Val Pro
405 410 415

Leu Glu Pro Arg Pro Leu Arg Arg Glu Ser Glu Ile
420 425

<210> SEQ ID NO 33
<211> LENGTH: 2332
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 33

```
gtctgcata gaa gagacca aaccaggcacc acttccttgc ttttccagc catgaatgct      60
tccagtcgga atgtgtttga cacgttgc acgggtgttga cagaaaatgtt gttcaaacat      120
cttcggaaat gggtcgtcac tcgctttttt gggcattctc ggcaaagagc aaggctagtc      180
tccaaagatg gaagggtgca a catagaattt ggcaatgtgg aggcacagtc aaggttata      240
ttctttgtgg acatctggac aacggtaactt gacctaagt ggagatacaa aatgaccatt      300
ttcatcacag cttcttggg gagttggttt ttctttggtc tcctgtggta tgcagttagcg      360
tacattcaca aagacacctcc ggaattccat cttctgcaca atcacactcc ctgtgtggag      420
```

-continued

aatattaatg	gcttgaccc	agctttctg	ttttctctgg	agactcaagt	gaccattgga	480
tatggattca	ggtgtgtgac	agaacagtgt	gccactgcca	ttttctgct	tatctttcag	540
tctataacttgc	gagttataat	caattcttc	atgtgtggg	ccatcttagc	caagatctcc	600
aggccccaaa	aacgtgccaa	gaccattacg	ttcagcaaga	acgcagtgat	cagcaaacgg	660
ggagggaaagc	tttgcctct	aatccgagtg	gctaattctca	ggaagagcct	tcttattggc	720
agtacattt	atggaaagct	tctgaagacc	acagtcactc	ctgaaggaga	gaccattatt	780
ttggaccaga	tcaatatcaa	ctttgttagtt	gacgctggg	atgaaaattt	attcttcattc	840
tccccattga	caatttacca	tgtcattgtat	cacaacagcc	ctttcttcca	catggcagcg	900
gagacccttc	tccagcagga	ctttgaattt	gtgggtttt	tagatggcac	agtggagtcc	960
accagtgcta	cctgccaagt	ccggacatcc	tatgtcccag	aggagggtgt	ttggggctac	1020
cgttttgctc	ccatagtatac	caagacaaag	gaagggaaat	accgagtgg	tttccataaac	1080
ttagcaaga	cagtggaaagt	ggagacccct	cactgtgcca	tgtgccttta	taatgagaaa	1140
gatgttagag	ccaggatgaa	gagaggctat	gacaacccca	acttcattt	gtcagaagtc	1200
aatgaaacag	atgacaccaa	aatgtAACAG	tggctttca	acgggagtaa	agcaaagtct	1260
ctaaagctcc	tagtacctag	aagcattatg	aagcagtcaa	caatttaggg	gtacgaaagt	1320
aggatgagag	ccttcaaagt	ctaccagcac	aaagacccct	gagccccca	attgtatcc	1380
cacaagacat	gcatctccac	aaggctactg	tattagaacg	tgcaatgcat	ttatataaaa	1440
ctgggtatg	gaagacatag	gtgctctctt	gaaatcttaa	atatgattat	ttgagctcat	1500
ataaggtgga	ttggagcaga	taaaattatc	aaaagttca	tgaacagggcc	aaacaaaataa	1560
tttttaaag	tttccttaaa	gaagttatga	actttagaaa	ggatcagggg	acaataataa	1620
tctcatttttgc	attctactga	taagaatgac	tccactttta	atgtggactt	ttactcatgg	1680
aaaaattgtc	tcctaatttgc	gggagatgaa	ccaaccaatc	aatgacaaga	aaacgcttac	1740
acaaagaaca	atttgaggct	ctaagcttct	catgtggtac	gttttagacag	aggctaaatc	1800
tgcacactag	aatcttgcgt	atacccttcc	gcaagacaga	atgctttagt	taaaagtgg	1860
gatgatattt	cattcaatct	gtattggatg	gcttaaaggg	ctataatct	gtttataaag	1920
agcatttccct	gctcttcgaa	gacagcaatg	aggagttgga	aggtgcaaag	tcagtagaga	1980
agggaaatgt	tcattaaatgc	acctgagaag	aaacagtttc	atgtgtttct	ccacctagag	2040
tttgcactgg	aatgtctattt	ctaaagaaga	agtggaaag	agagaggaat	gggatggagc	2100
cccacagtca	gaatgttact	atgtcttct	ttccctgaca	gcccatcttc	ctaaaaggaa	2160
ccagctttag	gaaggctcga	ccttgagggg	aaagtttac	tgtgaaagtc	ttcttcagat	2220
ccccacctgc	atcattccga	atgtgtctgt	gaaaaaaaaact	ggtactcaaa	gtctgtttagg	2280
aatcaaaatg	ttttcagtgt	gttgattat	atagtaaatt	tctgaaactg	tg	2332

<210> SEQ ID NO 34

<211> LENGTH: 391

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 34

Met	Asn	Ala	Ser	Ser	Arg	Asn	Val	Phe	Asp	Thr	Leu	Ile	Arg	Val	Leu
1							5						10		15

Thr Glu Ser Met Phe Lys His Leu Arg Lys Trp Val Val Thr Arg Phe

-continued

20	25	30
Phe	Gly	
His	Ser	Arg
35	36	37
Gln	Arg	Ala
40	41	42
Arg	Leu	Val
45	46	47
Lys	Asp	Gly
		Arg
Cys	Asn	Ile
50	51	52
Glu	Phe	Gly
		Asn
55	56	57
Val	Glu	Ala
		Gln
60	61	62
Ser	Arg	Phe
		Ile
65	66	67
Phe	Val	Asp
		Ile
70	71	72
Trp	Thr	Thr
		Val
75	76	77
Leu	Asp	Leu
		Gly
80	81	82
Lys	Trp	Arg
		Tyr
85	86	87
Met	Thr	Ile
		Phe
90	91	92
Ile	Thr	Ala
		Phe
95	96	97
Gly		
Leu	Leu	Trp
		Tyr
100	101	102
Ala	Val	Ala
		Tyr
105	106	107
Ile	His	Lys
		Asp
110	111	112
Leu	Pro	Glu
		Phe
115	116	117
His	Pro	Ser
		Ala
120	121	122
Asn	His	Thr
		Pro
125	126	127
Cys	Val	Glu
		Asn
130	131	132
Ala	Ser	Phe
		Ile
135	136	137
Gly	Leu	Glu
		Thr
140	141	142
Trp	Gly	Gln
		Cys
145	146	147
Phe	Arg	Cys
		Val
150	151	152
Thr	Glu	Gln
		Cys
155	156	157
Ala	Thr	Ala
		Ile
160	161	162
Phe	Ile	Phe
		Leu
165	166	167
Gly	Leu	Gly
		Val
170	171	172
Ile	Ile	Asn
		Ser
175	176	177
Phe	Gly	Met
		Cys
180	181	182
Ile	Ile	Asn
		Ser
185	186	187
Leu	Asp	Arg
		Val
190	191	192
Ala	Leu	Ala
		Lys
195	196	197
Thr	Phe	Ser
		Asn
200	201	202
Ala	Val	Ile
		Ser
205	206	207
Leu	Leu	Ile
		Arg
210	211	212
Arg	Val	Ala
		Asn
215	216	217
Leu	Arg	Lys
		Ser
220	221	222
Leu	Ile	Gly
		Ser
225	226	227
His	Ile	Tyr
		Gly
230	231	232
Leu	Leu	Lys
		Thr
235	236	237
Thr	Val	Thr
		Pro
240	241	242
Gly	Glu	Gly
		Leu
245	246	247
Thr	Ile	Ile
		Leu
250	251	252
Asp	Asp	Asp
		Gln
255	256	257
Ile	Ile	Ile
		Asn
260	261	262
Asn	Glu	Asn
		Leu
265	266	267
Phe	Ile	Ser
		Pro
270	271	272
Leu	Thr	Thr
		Ile
275	276	277
Asp	His	Asn
		Ser
280	281	282
Phe	Phe	His
		Met
285	286	287
His	Ala	Ala
		Glu
290	291	292
Gln	Asp	Phe
		Leu
295	296	297
Leu	Asp	Gly
		Thr
300	301	302
Val	Val	Val
		Phe
305	306	307
Ser	Ala	Thr
		Cys
310	311	312
Gln	Val	Arg
		Thr
315	316	317
Thr	Ser	Tyr
		Val
320	321	322
Glu	Pro	Glu
		Glu
325	326	327
Trp	Gly	Tyr
		Arg
330	331	332
Phe	Ala	Pro
		Ile
335	336	337
Asp	Val	Ser
		Lys
340	341	342
Tyr	Arg	Asp
		Phe
345	346	347
His	Asn	His
		Phe
350	351	352
Lys	Ser	Asp
		Val
355	356	357
Pro	His	Cys
		Ala
360	361	362
Met	Cys	Leu
		Tyr
365	366	367
Asn	Asn	Asn
		Glu
370	371	372
Met	Lys	Arg
		Gly
375	376	377
Tyr	Asp	Asp
		Thr
380	381	382
Lys	Asp	Asp
		Thr
385	386	387
Glu	Thr	Asp
		Asp
390	391	392
		Met

<210> SEQ ID NO 35

<211> LENGTH: 3102

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

-continued

<400> SEQUENCE: 35

tgttggtta	aaccacacaa	ctccacttt	gagttAACCA	ttgaaAGCTA	atgcaAGCAA	60
gcattcattgt	gaggcttaag	attcattaag	gtgggcctca	aagaagtCGG	cctttgcaca	120
gacaaaactg	aacagcacca	ctcacttgct	ttggccagca	tggatgttc	agatcgagg	180
tgtatgtgca	gagtgtgtat	cagggcactg	acagaaaAGGA	tgttcaaaca	tcttcgaaga	240
tggtttgtca	ctcacatatt	tggcggtct	cggcaacgag	caagggttgt	ctccaaagat	300
ggaagggtgt	acatcgagtt	tggcaatgt	gatgcacagt	cgagggttat	attcttgtg	360
gatatcttgg	caactgtact	tgacactgaaa	tggaggtaca	aatgaccgt	gttcattcaca	420
gccttcttgg	ggagttgggt	tctctttgtt	ctcctgttgt	atgtcgtagc	ctatgttcat	480
aaggatctcc	cagagtctca	cccacctgac	aaccgtactc	cttgtgtgga	gaacattaa	540
ggcatgacat	cagectttct	gtttctcta	gagacccaag	tgaccatagg	ttacggattc	600
aggtttgtga	cagaacagt	tgccactgcc	atttttctgc	ttatcttcca	gtctatttt	660
ggagtgtatca	tcaattcttt	catgtgttgt	gccatattag	ccaagatctc	tagacccaa	720
aaacgtgcaa	agaccattac	attcagcaag	aatgcgggt	tcagcaaaacg	tggggggaa	780
ctctgtctcc	tcatccgagt	agcaaatttt	aggaaaagcc	ttctgattgg	cagtcacata	840
tatggtaagc	ttctgaagac	taccatcaca	cctgaaggag	agaccattat	tttggatcag	900
accaatataa	actttgttgt	tgtgctggc	aatgaaaatt	tgttcttcat	ttccccactg	960
acaatctacc	acattattga	ccacaacagc	cctttcttcc	acatggcggc	agaaaacttt	1020
tcccaacagg	acttcgagtt	ggttgtcttt	ttagatggca	cagtagaaatc	caccagtgc	1080
acctgccaag	tccgcacatc	atacatccca	gaagaggtgc	tttggggta	ccgttttgtt	1140
cccatcgat	ccaagaccaa	ggaaggggaaa	taccgagttg	atttccataa	ctttgttaag	1200
acgggtggaa	tggagacccc	tcatgtgc	atgtgcctct	ataatgagaa	agatgcagg	1260
gccaggatga	agagaggcgt	tgacaaccc	aactttgtct	tgtcagaagt	tgatgaaaca	1320
gacgacaccc	aaatgttagc	gtggctttc	tacctacaaa	gagtctctca	aggacctaag	1380
tgctggctgt	gttcagaagc	atcctgcagt	tggggactg	aaagcaggat	gagaacccaa	1440
gaagtctacc	agcagagtca	ccccctcagc	cccatggct	tgattctaga	agactcgtag	1500
ctctaaagtc	tgcagcagta	agcatacggt	ggaagcatgt	gaatgaacct	ggcgtgtgga	1560
agccaaaagg	agtcacttg	gatcttca	ctgactgtgt	aagctcgatc	agtgttgatg	1620
gaaacaaagt	catccatga	catgtactgg	agaaaaagt	ggagcttag	aaaacttcag	1680
gagctagcca	tatttcctgt	ttgattctat	ggatgagaaa	gatatcacat	tttatcttaa	1740
agttagactt	tatcaacgg	aaatctgtct	tctgagttgg	gaagtggaccc	aaccagtca	1800
tgacaataag	agectgtcat	acacagaatc	actaaaggct	ctaaacttctt	aagctctgct	1860
gtttgaagcc	tttgcgtct	tctgggtcca	tgtttcagaa	ggggtaaggt	gggcattccag	1920
tgactgcacc	tttctcaacc	caaggtacag	aagaacagaa	agcctcaacc	aatttataa	1980
ctaaactcaga	cgctgcagcc	catataggaa	tggacctgaa	tgattccgg	tgggtcaatc	2040
aatggaggcc	aagcccatc	cttagactaa	atgaacatc	tcttgcaaa	gagagtgttt	2100
gagctaaaag	tgttgggtaa	gaggctgaag	agatgcatt	ctttgttaag	tcatgtgcg	2160
cacgcgtgaa	gacctgggtc	tgaacccatc	acaatcacat	acaaggctt	acttagtagg	2220

-continued

atgctccccc tccccaaatcc tagtactggg gaggtaggga taggaggatg tctcaagttt	2280
tctggacago tgaatgggtg aggtaccagt tcagggagag actacaaaaa taaggtgaag	2340
gggctggcgg gaagactctg gttaaaagca ccgtggctga tcttccagag aaccttagtt	2400
tggttcctgg cacagttactt cataactatc tataactcca gtttcagggc atctgtatgc	2460
cttttctggc ccctatgagc accaggtgca catgcgagca aaacactata cacatacgat	2520
taaaaagata aggttagagaa tgactgaaga agaaatctta cgtcaacttc tgatctacac	2580
acacacacac acacacacac acacacacac acagagcata gtgatgttg tttcagttt	2640
cagtgggtag gttataggc cacaggtctg gtaaaaagagg ttttattctt tcctctctca	2700
agatggtgct gaggactggg aaagcagtca gcctggagtg tgtatgtgc taacaccgaa	2760
tggttcttgtt gtttctctga ctgtggttg aactgcaatt cttaggtatga aggagaaaagg	2820
aacgagatac agcaccacag tcaggatgct gctgctttct tgtctctaaa ggccatccct	2880
acatcttcca aagggaaacaa gttagcagga gaatttactt cgaggaaagg ttctaattgtc	2940
ttgtttgttc atttccacat ctgcatttcatt ctgatttcac cttgggggaa aaaaagttaga	3000
attcaaagct gttcaaaaat caaaatatca ttttagtatg tgattaataa atttctgaaa	3060
ctgtatgagc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa	3102

<210> SEQ_ID NO 36

<211> LENGTH: 372

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 36

Met Phe Lys His Leu Arg Arg Trp Phe Val Thr His Ile Phe Gly Arg	
1 5 10 15	
Ser Arg Gln Arg Ala Arg Leu Val Ser Lys Asp Gly Arg Cys Asn Ile	
20 25 30	
Glu Phe Gly Asn Val Asp Ala Gln Ser Arg Phe Ile Phe Phe Val Asp	
35 40 45	
Ile Trp Thr Thr Val Leu Asp Leu Lys Trp Arg Tyr Lys Met Thr Val	
50 55 60	
Phe Ile Thr Ala Phe Leu Gly Ser Trp Phe Leu Phe Gly Leu Leu Trp	
65 70 75 80	
Tyr Val Val Ala Tyr Val His Lys Asp Leu Pro Glu Phe Tyr Pro Pro	
85 90 95	
Asp Asn Arg Thr Pro Cys Val Glu Asn Ile Asn Gly Met Thr Ser Ala	
100 105 110	
Phe Leu Phe Ser Leu Glu Thr Gln Val Thr Ile Gly Tyr Gly Phe Arg	
115 120 125	
Phe Val Thr Glu Gln Cys Ala Thr Ala Ile Phe Leu Leu Ile Phe Gln	
130 135 140	
Ser Ile Leu Gly Val Ile Ile Asn Ser Phe Met Cys Gly Ala Ile Leu	
145 150 155 160	
Ala Lys Ile Ser Arg Pro Lys Lys Arg Ala Lys Thr Ile Thr Phe Ser	
165 170 175	
Lys Asn Ala Val Ile Ser Lys Arg Gly Gly Lys Leu Cys Leu Leu Ile	
180 185 190	
Arg Val Ala Asn Leu Arg Lys Ser Leu Leu Ile Gly Ser His Ile Tyr	

-continued

195	200	205	
Gly Lys Leu Leu Lys Thr Thr Ile Thr Pro Glu Gly Glu Thr Ile Ile			
210	215	220	
Leu Asp Gln Thr Asn Ile Asn Phe Val Val Asp Ala Gly Asn Glu Asn			
225	230	235	240
Leu Phe Phe Ile Ser Pro Leu Thr Ile Tyr His Ile Ile Asp His Asn			
245	250	255	
Ser Pro Phe Phe His Met Ala Ala Glu Thr Leu Ser Gln Gln Asp Phe			
260	265	270	
Glu Leu Val Val Phe Leu Asp Gly Thr Val Glu Ser Thr Ser Ala Thr			
275	280	285	
Cys Gln Val Arg Thr Ser Tyr Ile Pro Glu Glu Val Leu Trp Gly Tyr			
290	295	300	
Arg Phe Val Pro Ile Val Ser Lys Thr Lys Glu Gly Lys Tyr Arg Val			
305	310	315	320
Asp Phe His Asn Phe Gly Lys Thr Val Glu Val Glu Thr Pro His Cys			
325	330	335	
Ala Met Cys Leu Tyr Asn Glu Lys Asp Ala Arg Ala Arg Met Lys Arg			
340	345	350	
Gly Tyr Asp Asn Pro Asn Phe Val Leu Ser Glu Val Asp Glu Thr Asp			
355	360	365	
Asp Thr Gln Met			
370			

<210> SEQ_ID NO 37
 <211> LENGTH: 4187
 <212> TYPE: DNA
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 37

ggcggcgca	gggcggggct	ttacggacgc	aagcacgtcg	aagcgctgct	cctggagccg	60
cggagggtgc	gggttggct	gccccgttgtt	ctgtggcggt	tgctgtggcg	gagtttggag	120
gttggagaga	aatccaggtta	ctcaactagac	tgg taccttc	tgccaccatg	ggggagctt	180
tccggagtga	agaaatgaca	ctggcccagc	tttttctaca	gtcagaggct	gcttattgtt	240
gtgtcagtga	attaggagaa	cttggaaagg	ttcagtttcg	tgacttaaat	ccagatgtga	300
atgttttcca	acggaaattt	gtgaatgaag	ttagaagatg	tgaagaaatg	gatcgaaagc	360
ttcgatttgt	tgagaaaagag	ataagaaaag	ctaacattcc	gattatggac	accggtaaaa	420
acccagaggt	tcccttcccc	cgggacatga	ttgacttaga	ggccaatttt	gagaagattg	480
aaaatgaact	gaaggaaatc	aacacaaacc	aggaagctct	gaagagaaac	tccctggAAC	540
tgaccgaatt	aaaatttata	cttcgcaaaa	ctcagcaatt	ttttgatgag	atggcgatc	600
cagactgtt	ggaagagtcc	tcatccctct	tggagccaag	tgagatggga	agaggtactc	660
ctttaagact	tggcttcgtg	gctgggtgtca	ttaaccggga	gccccatccct	acttttggac	720
gcatgcttg	cgccggatgc	cgccggaaatg	tgttccctgcg	acaggctgaa	atcgagaacc	780
ccctggagga	tcctgtgact	ggcgactacg	tgcacaagtc	tgtgtttatc	attttcttcc	840
aaggcgatca	gctaaaaaac	agagtcaaga	aaatctgtga	agggttccga	gcctcactct	900
atccctgtcc	tgagacacca	caggagagga	aggaaatggc	ttctggagtg	aataccagga	960
ttgtatgatct	ccaaatggtt	ctgaatcaa	cgaggatca	ccgcccagagg	gttctgcagg	1020

-continued

cagctgctaa	gaacatccgt	gtctggttca	tcaaagtgcg	gaagatgaag	gccatctatc	1080
acaccctgaa	cctgtgcaac	atagatgtga	ctcagaaatg	cttgattgca	gagggtctgg	1140
gccctgtcac	cgacccttgcac	tccatccagt	ttgcactcag	aaggggcacg	gaacacagtg	1200
gttccactgt	accttccatt	ttgaacagga	tgcagacaaa	ccagactccc	ccaaacctata	1260
acaaaaacaa	caagtttacc	tatggcttc	agaacatagt	agatgcttat	ggaattggaa	1320
cttaccgaga	gataaatcca	gctccgtata	ctattatcac	gttccctttt	ctatttgctg	1380
tgtatgtttgg	agacttcggt	catggcattt	taatgaccct	ttttgctgtg	tggatggtag	1440
tgagggagag	ccggatcctt	tcccagaaga	atgagaatga	gatgtttagc	actgtgttca	1500
gtggtcgata	cattattta	ttgatgggtg	tgttctccat	gtacactggc	ctcatctaca	1560
atgattgctt	ttccaagtct	cttaatatct	ttgggtcattc	ctggagtgta	cgcccgatgt	1620
ttacttataa	ttggactgaa	gagacgcttc	gggggaaccc	tgttctacag	ctgaacccag	1680
ccctccctgg	agtgtttgg	ggaccatacc	cttttggcat	tgttccattt	tggaaacattg	1740
ctaccaataa	actgacgttc	ttgaactctt	ttaagatgaa	gatgtctgtt	atccttggta	1800
tcatccat	gctgtttgga	gtcagcgtga	gtctgttcaa	ccatcatctat	ttcaagaagc	1860
ccctgaatat	ctactttgga	tttattctcg	aaataatctt	catgacccct	ttgtttggct	1920
atttggttat	ccttattttt	tacaagtggaa	cgccctatga	tgctcataacc	tctgagaatg	1980
caccaagcct	tctgatccat	ttcataaaaca	tgttcccttt	ttcctaccca	gagtctgggt	2040
atccaatgtt	gtattctgga	cagaaaggaa	ttcagtgttt	cctggtagtg	gttgcactac	2100
tgtgtgtacc	ttggatgtcg	ctgtttaaac	cattggtcct	tcggcgtag	tatggagga	2160
gaaagcattt	gggaactctc	aactttgggt	ggatcagggt	ggcaacggg	ccgacagagg	2220
aggatgctga	gattattcag	catgaccaggc	tctccaccca	ctcagaggac	gcagacaggt	2280
ttgactttgg	ggacaccatg	gtccaccagg	ccatccacac	catcgagtag	tgcctgggt	2340
gcacatccaa	caactgcctcc	tacttgeggc	tctggccct	cagcctcgct	catgegcage	2400
tgtctgaggt	gctttggacc	atgggtatcc	acatggcct	gagcgtgaag	agcttggcg	2460
gaggtttgg	gctgttcttc	tttttcactg	cctttgeccac	cctgaccgtg	gccatctcc	2520
tgtatcatgga	gggcctctcg	gcctttctcc	acgcactgctg	cttacactgg	gtttagttcc	2580
agaataaaatt	ctacagcggg	accggttca	agttcttacc	cttctcttc	gagcatattc	2640
gggaaggggaa	gtttgaagag	ttagtccctg	ttagggccgt	gtgccccatg	ctaccctccc	2700
cgcctccctc	cacagtgtatc	agctgtgcct	ctctgcctgt	tggtgtgtat	ctgtggcgac	2760
cagetcatcc	gtgtcacect	gtctgtgagt	catttagata	gaatagtct	ccttgggtct	2820
cccaccaccc	ctagcttgc	gtgttagtgta	gtgattttct	ggctgtcact	cataactcact	2880
ggcaccaggc	cttgcctct	tagcctccat	ccatccagac	agcccttccc	acctcttgg	2940
ggtgagccag	tctgcattcc	cacgccatcc	caaagccctt	tcatctccc	cgtgcattgt	3000
agatggaagg	agcacccatg	ccattcacat	ctagactttg	agttccctgc	atctgcacc	3060
gtatgtttcta	gcaggagtag	tggggggagt	aatacagatt	cttccctaga	agggggacact	3120
gttaacatgt	cccactcttg	gattagcagg	ggtgggtcca	ggaagatgtat	atttgcgtct	3180
tttgcctacc	ccccctggcat	tcagctggac	ccaaacttaggc	catcatgagt	ggcttctccc	3240
tgtcatcccc	aggggtcata	ggatatactac	accgccttcc	tgacccacc	ctgcactccc	3300

-continued

atccttcct ctctccccgt tcatgcctg cactacatag cacagccggg atgcttggaa	3360
cagaggcctt ggctgctcg cagtgcacag ggcttccctc tctcggggtt ggcttcttcc	3420
caggccttgc atggccctg cccacaagca caccctcagg ccgagggtgc agactgatgc	3480
tcttccctga tggagaccct gagatcttcc ccaccccaa tcatgatgtc ttcaagtgtgg	3540
gactggggtc ctcttggttc tgccctgcagg ctgcctggct ccgccttag tgcccctcc	3600
tcaccacact ggccccaggt ctcaggagggt gtgtcctggg cagggaaaggt cagtgtcact	3660
gatggtttgc tgtttggaaag ccattggcag ggctgcccgt catgtggctg tgagggtgc	3720
acagtcctgc caaggggctt ctccttgctc accccgaaacc ttgtaatcgt gtgctggcgt	3780
ggcagccctg gctaagttaa tccccaccgc tttcagtggt agaaagaatt ccctgagtgg	3840
gccaggctgg tgcccttcctc ctaccctggc ttttctgagt gagctgcctg gagccctcat	3900
ccctctccc aggctgggtt ggccctgggc gggccactg tgtgctggcc cactgtgacc	3960
tgacccgacc ttgtgcagcc cccctgcctt ggtgtcctgg gttttctgta tgatcttgc	4020
tctgtttcca gtggggtttg aagcagagtt cagggaaaccc tgcccaaggt ctcctgttc	4080
agacattcct atgttgaata aagtatgttt gacttccccc gaaaaaaaaaaaaaaa	4140
aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa	4187

<210> SEQ_ID NO 38

<211> LENGTH: 831

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 38

Met Gly Leu Phe Arg Ser Glu Glu Met Thr Leu Ala Gln Leu Phe	
1 5 10 15	
Leu Gln Ser Glu Ala Ala Tyr Cys Cys Val Ser Glu Leu Gly Glu Leu	
20 25 30	
Gly Lys Val Gln Phe Arg Asp Leu Asn Pro Asp Val Asn Val Phe Gln	
35 40 45	
Arg Lys Phe Val Asn Glu Val Arg Arg Cys Glu Glu Met Asp Arg Lys	
50 55 60	
Leu Arg Phe Val Glu Lys Glu Ile Arg Lys Ala Asn Ile Pro Ile Met	
65 70 75 80	
Asp Thr Gly Glu Asn Pro Glu Val Pro Phe Pro Arg Asp Met Ile Asp	
85 90 95	
Leu Glu Ala Asn Phe Glu Lys Ile Glu Asn Glu Leu Lys Glu Ile Asn	
100 105 110	
Thr Asn Gln Glu Ala Leu Lys Arg Asn Phe Leu Glu Leu Thr Glu Leu	
115 120 125	
Lys Phe Ile Leu Arg Lys Thr Gln Gln Phe Phe Asp Glu Met Ala Asp	
130 135 140	
Pro Asp Leu Leu Glu Glu Ser Ser Ser Leu Leu Glu Pro Ser Glu Met	
145 150 155 160	
Gly Arg Gly Thr Pro Leu Arg Leu Gly Phe Val Ala Gly Val Ile Asn	
165 170 175	
Arg Glu Arg Ile Pro Thr Phe Glu Arg Met Leu Trp Arg Val Cys Arg	
180 185 190	
Gly Asn Val Phe Leu Arg Gln Ala Glu Ile Glu Asn Pro Leu Glu Asp	

-continued

195	200	205
Pro Val Thr Gly Asp Tyr Val His Lys Ser Val Phe Ile Ile Phe Phe		
210	215	220
Gln Gly Asp Gln Leu Lys Asn Arg Val Lys Lys Ile Cys Glu Gly Phe		
225	230	235
Arg Ala Ser Leu Tyr Pro Cys Pro Glu Thr Pro Gln Glu Arg Lys Glu		
245	250	255
Met Ala Ser Gly Val Asn Thr Arg Ile Asp Asp Leu Gln Met Val Leu		
260	265	270
Asn Gln Thr Glu Asp His Arg Gln Arg Val Leu Gln Ala Ala Ala Lys		
275	280	285
Asn Ile Arg Val Trp Phe Ile Lys Val Arg Lys Met Lys Ala Ile Tyr		
290	295	300
His Thr Leu Asn Leu Cys Asn Ile Asp Val Thr Gln Lys Cys Leu Ile		
305	310	315
Ala Glu Val Trp Cys Pro Val Thr Asp Leu Asp Ser Ile Gln Phe Ala		
325	330	335
Leu Arg Arg Gly Thr Glu His Ser Gly Ser Thr Val Pro Ser Ile Leu		
340	345	350
Asn Arg Met Gln Thr Asn Gln Thr Pro Pro Thr Tyr Asn Lys Thr Asn		
355	360	365
Lys Phe Thr Tyr Gly Phe Gln Asn Ile Val Asp Ala Tyr Gly Ile Gly		
370	375	380
Thr Tyr Arg Glu Ile Asn Pro Ala Pro Tyr Thr Ile Ile Thr Phe Pro		
385	390	395
Phe Leu Phe Ala Val Met Phe Gly Asp Phe Gly His Gly Ile Leu Met		
405	410	415
Thr Leu Phe Ala Val Trp Met Val Leu Arg Glu Ser Arg Ile Leu Ser		
420	425	430
Gln Lys Asn Glu Asn Glu Met Phe Ser Thr Val Phe Ser Gly Arg Tyr		
435	440	445
Ile Ile Leu Leu Met Gly Val Phe Ser Met Tyr Thr Gly Leu Ile Tyr		
450	455	460
Asn Asp Cys Phe Ser Lys Ser Leu Asn Ile Phe Gly Ser Ser Trp Ser		
465	470	475
480		
Val Arg Pro Met Phe Thr Tyr Asn Trp Thr Glu Glu Thr Leu Arg Gly		
485	490	495
Asn Pro Val Leu Gln Leu Asn Pro Ala Leu Pro Gly Val Phe Gly Gly		
500	505	510
Pro Tyr Pro Phe Gly Ile Asp Pro Ile Trp Asn Ile Ala Thr Asn Lys		
515	520	525
Leu Thr Phe Leu Asn Ser Phe Lys Met Lys Met Ser Val Ile Leu Gly		
530	535	540
Ile Ile His Met Leu Phe Gly Val Ser Leu Ser Leu Phe Asn His Ile		
545	550	555
560		
Tyr Phe Lys Lys Pro Leu Asn Ile Tyr Phe Gly Phe Ile Pro Glu Ile		
565	570	575
Ile Phe Met Thr Ser Leu Phe Gly Tyr Leu Val Ile Leu Ile Phe Tyr		
580	585	590
Lys Trp Thr Ala Tyr Asp Ala His Thr Ser Glu Asn Ala Pro Ser Leu		
595	600	605

-continued

Leu Ile His Phe Ile Asn Met Phe Leu Phe Ser Tyr Pro Glu Ser Gly
 610 615 620
 Tyr Ser Met Leu Tyr Ser Gly Gln Lys Gly Ile Gln Cys Phe Leu Val
 625 630 635 640
 Val Val Ala Leu Leu Cys Val Pro Trp Met Leu Leu Phe Lys Pro Leu
 645 650 655
 Val Leu Arg Arg Gln Tyr Leu Arg Arg Lys His Leu Gly Thr Leu Asn
 660 665 670
 Phe Gly Gly Ile Arg Val Gly Asn Gly Pro Thr Glu Glu Asp Ala Glu
 675 680 685
 Ile Ile Gln His Asp Gln Leu Ser Thr His Ser Glu Asp Ala Asp Glu
 690 695 700
 Phe Asp Phe Gly Asp Thr Met Val His Gln Ala Ile His Thr Ile Glu
 705 710 715 720
 Tyr Cys Leu Gly Cys Ile Ser Asn Thr Ala Ser Tyr Leu Arg Leu Trp
 725 730 735
 Ala Leu Ser Leu Ala His Ala Gln Leu Ser Glu Val Leu Trp Thr Met
 740 745 750
 Val Ile His Ile Gly Leu Ser Val Lys Ser Leu Ala Gly Leu Val
 755 760 765
 Leu Phe Phe Phe Thr Ala Phe Ala Thr Leu Thr Val Ala Ile Leu
 770 775 780
 Leu Ile Met Glu Gly Leu Ser Ala Phe Leu His Ala Leu Arg Leu His
 785 790 795 800
 Trp Val Glu Phe Gln Asn Lys Phe Tyr Ser Gly Thr Gly Phe Lys Phe
 805 810 815
 Leu Pro Phe Ser Phe Glu His Ile Arg Glu Gly Lys Phe Glu Glu
 820 825 830

<210> SEQ ID NO 39

<211> LENGTH: 4516

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 39

```

ctcggcgccgc ctggaccctt gcccctctctt gggtgagaa gtcggggcc gttccgggt      60
ttcactcctt ctcagcctgg gtcggccagcc cctctctctt tttctggac tggctctcac      120
ccccttcggt ccccttcctt tagctcaggc tccctacccc ttcccttagc ccacagccca      180
gagtccccagc tcctcagtc cttccctcag ccaaagggtcc cagccttcct tcttccttc      240
ctttgcacta tccctatactt gccccttcctt ctatccctag ggctcagttt cccacatccg      300
tcctcccccctt tccctaggccc ggagttccag accttttgtt ctccttcgtt ggctgttctt      360
gggtcccttc cccctttccc cactttggat ttccagatgtt caaacccagc ctccctccac      420
ccccagaaaa ttgtttccat ggaaatgcct ctctaaaaca tgaacttttc cttagagacta      480
cgccagtc tcttccactt tgctgaccctt ttgctaccta tggccgggtt ttactctca      540
tttggtaag gtcgaggctg gctctggaaag cagcaccatg gttctgggtt ctggcatctg      600
tggcctctt ccacatcgga tcttcccttc cttactcggtt gtgggttctt tggtggggct      660
gctgcctgtt ctcaggagcc atggcctcca gctcagccca actgccagca ccattcgaag      720
ctcagagccca ccacgagaac gctcgattgg ggatgtcacc accgctccac cggaggtcac      780

```

-continued

cccgagagc	cggccgttta	atcattccgt	cactgatcat	ggcatgaagc	cgcgcaaggc	840
ctttccagtc	ctggggcatcg	actacacaca	cgtgcgcacc	cccttcgaga	tctccctctg	900
gatccttctg	gcctgcctca	tgaagatagg	tttccatgtg	atccccacta	tctcaagcat	960
cgtcccggag	agctgcctgc	tgatcgtgg	ggggctgctg	gtggggggcc	tgatcaaggg	1020
tgttaggcgag	acacccccc	tccgcagtc	cgacgtcttc	ttcctcttcc	tgctgcgc	1080
catcatcctg	gatgcgggct	acttcctgcc	actgcggcag	ttcacagaaa	acctgggcac	1140
catcctgatc	tttgcgtgg	tgggcacgct	gtggAACGCC	ttcttcctgg	gcggcctcat	1200
gtacgcgtg	tgccctgggg	gccccgtggaa	gatcaacaac	atcggccctcc	tggacaacct	1260
gctttcggc	agcatcatct	cggccgtgga	ccccgtggcg	gttctggctg	tctttgagga	1320
aattcacatc	aatgagctgc	tgcacatct	tgtttttggg	gagtccctgc	tcaatgacgc	1380
cgtcaactgtg	gtccctgtatc	acctcttga	ggagtttgc	aactacgaac	acgtggccat	1440
cgtggacatc	ttcctcggct	tcctgagctt	cttcgtggtg	gcccctggcg	gggtgtttgt	1500
gggggtggtc	taaaaaaaaa	tcgcagccct	ccccccccca	tttacctccc	acatccgggt	1560
catcgagccg	cttttcgtct	tcctctacag	ctacatggcc	tacttgtcag	ccgagctctt	1620
ccacctgtca	ggcatcatgg	cgctcatagc	ctcaggagtg	gtgatgcgc	cctatgtgga	1680
ggccaacatc	tcccaacaatg	ccccacaccac	catcaaatac	ttccctgaaga	tgtggaggag	1740
cgtcagcgag	accctcatct	tcatcttcc	cggcgtctcc	acgggtggcg	gtccccacca	1800
ctggaaactgg	acccctcgta	tcagcaccc	gtctttctgc	ctcatcgccc	gctgtgtggg	1860
ggtgctgggc	ctgacctgg	tcatcaacaa	gttccgtatc	gtgaagctga	cccccaagga	1920
ccagttcatac	atcgccatag	ggggcctgc	agggggccatc	gccttctctc	tgggetacct	1980
cctggacaag	aaggacttcc	ccatgtgtga	cctgttcc	actgcacatca	tcactgtcat	2040
cttttcacc	gtctttgtgc	agggcatgac	cattccccc	ctggtagacc	tgttggctgt	2100
gaagaaaaaa	caagagacga	agcgctccat	caacgaagag	atccacacac	agtccctgga	2160
ccaccttctg	acaggcatcg	aagacatctg	tggccactac	ggtcaccacc	actggaaagg	2220
caagctcaac	cggtttaata	agaaatatgt	gaagaagtgt	ctgatagctg	gcgagcgctc	2280
caaggagccc	cagctcattt	cattctacca	caagatggag	atgaaggcagg	ccatcgagct	2340
ggtgagagac	gggggcattgg	gcaagatccc	ctctgcgtc	tccacccgt	ccatgcagaa	2400
catccacccc	aagtccctgc	cttcgcagcg	catcctgc	gcactgtcca	aggacaagg	2460
ggaggagatc	cgcggaaatcc	tgaggaacaa	cttgcagaag	accaggcagc	ggctgcgggc	2520
ctacaacaga	cacacgctgg	tggcagaccc	ctacgaggg	gcctggaaacc	agatgtgt	2580
ccggaggcag	aaggccccggc	agctggagca	gaagatcaac	aactacctga	cggtgcgc	2640
ccacaagctg	gactcaccca	ccatgtctcg	ggcccgc	ggctcagacc	cactggct	2700
tgagccgaag	gaggacctgc	ctgtcatcac	catcgaccc	gttccccgc	agtccccc	2760
gtctgtggac	ctgggtgaatg	aggagctgaa	ggccaaagtc	ttagggtga	gcccggatcc	2820
tgcaaagggt	gctgaggagg	acgaggacga	cgtgggggc	atcatgtgc	ggagcaagg	2880
gacttcgtcc	ccaggaaccg	acgatgtctt	caccccccgc	cccagtgaca	gccccagctc	2940
ccagaggata	cagcgctgcc	tcagtgaccc	aggcccacac	cctgagctg	ggggaggaga	3000
accgttctc	cccaaggggc	agtaacgcca	gggcacgcag	gcagcgctg	tccctcaca	3060

-continued

gactttcca ccagagcagg ggctgctggg ggctccctt gcccctcctg acccgattg	3120
gccctgcccc tccccctacc gcatggcgc tggggccaca gccccccaccc cagcacagct	3180
cctccccctgc cgccctccgg gaagcatcct ccccaccaga gctgcctccc caatccattt	3240
ggcagaactg ctggggctgg tgaggccggc cctgccccctc cctagatcca ggcttcctcc	3300
ggacctggac tagggcctcg gaggctcctc cctctgcctc atcctccctc tcattcagac	3360
caatcttagt ttctaaccaa agagtctctg gctcagctgt ggtcccccaccc aggaaggggag	3420
ggagctgagg cctcccttga gtaggccctg ctttatcagg ggacaaacca ggggtaccag	3480
gcacatggc gggggaggga ctgctgaccc accaaggctct cacactccctc ctgcagctc	3540
tgtcaccctg gccaccaccc aacctatcct tactcagagc tgccggctga gggcatctct	3600
gagtgtctct gcctggagca ggggtggttt ctacggtgac agtgcgtga ctcagagctt	3660
ttcgaactgt gctcccacgg ggaccactgg gcccctcagg ggaagctgt aggggaagga	3720
ctggccgtgg ctccagaatg tgctgcctt ttaagtttg tttgttcaca ctcctata	3780
tgattgtttg cacagagggc gtcctgttt ttaaaaacatt ttgaaaacccc ctggctgaac	3840
agtgcctctgc ctcttaactcc ctccctcacac tccagaatata cccttcctca tctgtgcctg	3900
tctgtccaac ccctccccca cgtctctctg cctgctggc tcttaactgt tgctcgaaga	3960
ctgtgacatc agaagtaact cccactcota atcaagagtc tctccagcct cacagatgt	4020
ggcctcttgg cacctgccta gtcctgggc ctgaccctcca gtcctgtgg cctgtctta	4080
cttcccccac cctgggtttt gcccctggaa ctttccctt gtgtgtacca caccctgcct	4140
gctgtggagc ccattgtgga ggccgtgggg gggagaaggc ctcctctgag gatccctgt	4200
ccctggggc ttgtggattt ggcagaatcc tggggcccca gagaccttgc cccacacaca	4260
ctccttcccc ttgtccctgg ggcactcccc caggattgtg caatagtcag agtgcctt	4320
tttgcaggggg actggggcat gggtcctcgcc cccatctgtc catccctcctc tccatgcag	4380
tgctgtttgg gcaggagtca ccatgcaagg gtgacatcga caaccacgta ccaagccacc	4440
gcagctgctg ccactctgtc gcctgtacag aagaaactga atcttttca tattctaata	4500
aatcaatgtg agtttt	4516

<210> SEQ ID NO 40

<211> LENGTH: 815

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40

Met Val Leu Arg Ser Gly Ile Cys Gly Leu Ser Pro His Arg Ile Phe			
1	5	10	15

Pro Ser Leu Leu Val Val Ala Leu Val Gly Leu Leu Pro Val Leu			
20	25	30	

Arg Ser His Gly Leu Gln Leu Ser Pro Thr Ala Ser Thr Ile Arg Ser			
35	40	45	

Ser Glu Pro Pro Arg Glu Arg Ser Ile Gly Asp Val Thr Thr Ala Pro			
50	55	60	

Pro Glu Val Thr Pro Glu Ser Arg Pro Val Asn His Ser Val Thr Asp			
65	70	75	80

His Gly Met Lys Pro Arg Lys Ala Phe Pro Val Leu Gly Ile Asp Tyr			
85	90	95	

-continued

Thr His Val Arg Thr Pro Phe Glu Ile Ser Leu Trp Ile Leu Leu Ala
 100 105 110
 Cys Leu Met Lys Ile Gly Phe His Val Ile Pro Thr Ile Ser Ser Ile
 115 120 125
 Val Pro Glu Ser Cys Leu Leu Ile Val Val Gly Leu Leu Val Gly Gly
 130 135 140
 Leu Ile Lys Gly Val Gly Glu Thr Pro Pro Phe Leu Gln Ser Asp Val
 145 150 155 160
 Phe Phe Leu Phe Leu Leu Pro Pro Ile Ile Leu Asp Ala Gly Tyr Phe
 165 170 175
 Leu Pro Leu Arg Gln Phe Thr Glu Asn Leu Gly Thr Ile Leu Ile Phe
 180 185 190
 Ala Val Val Gly Thr Leu Trp Asn Ala Phe Phe Leu Gly Leu Met
 195 200 205
 Tyr Ala Val Cys Leu Val Gly Gly Glu Gln Ile Asn Asn Ile Gly Leu
 210 215 220
 Leu Asp Asn Leu Leu Phe Gly Ser Ile Ile Ser Ala Val Asp Pro Val
 225 230 235 240
 Ala Val Leu Ala Val Phe Glu Glu Ile His Ile Asn Glu Leu Leu His
 245 250 255
 Ile Leu Val Phe Gly Glu Ser Leu Leu Asn Asp Ala Val Thr Val Val
 260 265 270
 Leu Tyr His Leu Phe Glu Glu Phe Ala Asn Tyr Glu His Val Gly Ile
 275 280 285
 Val Asp Ile Phe Leu Gly Phe Leu Ser Phe Phe Val Val Ala Leu Gly
 290 295 300
 Gly Val Leu Val Gly Val Val Tyr Gly Val Ile Ala Ala Phe Thr Ser
 305 310 315 320
 Arg Phe Thr Ser His Ile Arg Val Ile Glu Pro Leu Phe Val Phe Leu
 325 330 335
 Tyr Ser Tyr Met Ala Tyr Leu Ser Ala Glu Leu Phe His Leu Ser Gly
 340 345 350
 Ile Met Ala Leu Ile Ala Ser Gly Val Val Met Arg Pro Tyr Val Glu
 355 360 365
 Ala Asn Ile Ser His Lys Ser His Thr Thr Ile Lys Tyr Phe Leu Lys
 370 375 380
 Met Trp Ser Ser Val Ser Glu Thr Leu Ile Phe Ile Phe Leu Gly Val
 385 390 395 400
 Ser Thr Val Ala Gly Ser His His Trp Asn Trp Thr Phe Val Ile Ser
 405 410 415
 Thr Leu Leu Phe Cys Leu Ile Ala Arg Val Leu Gly Val Leu Gly Leu
 420 425 430
 Thr Trp Phe Ile Asn Lys Phe Arg Ile Val Lys Leu Thr Pro Lys Asp
 435 440 445
 Gln Phe Ile Ile Ala Tyr Gly Gly Leu Arg Gly Ala Ile Ala Phe Ser
 450 455 460
 Leu Gly Tyr Leu Leu Asp Lys Lys His Phe Pro Met Cys Asp Leu Phe
 465 470 475 480
 Leu Thr Ala Ile Ile Thr Val Ile Phe Phe Thr Val Phe Val Gln Gly
 485 490 495

-continued

Met	Thr	Ile	Arg	Pro	Leu	Val	Asp	Leu	Leu	Ala	Val	Lys	Lys	Lys	Gln
500					505							510			
Glu	Thr	Lys	Arg	Ser	Ile	Asn	Glu	Glu	Ile	His	Thr	Gln	Phe	Leu	Asp
515					520						525				
His	Leu	Leu	Thr	Gly	Ile	Glu	Asp	Ile	Cys	Gly	His	Tyr	Gly	His	His
530					535					540					
His	Trp	Lys	Asp	Lys	Leu	Asn	Arg	Phe	Asn	Lys	Lys	Tyr	Val	Lys	Lys
545					550					555			560		
Cys	Leu	Ile	Ala	Gly	Glu	Arg	Ser	Lys	Glu	Pro	Gln	Leu	Ile	Ala	Phe
565					570							575			
Tyr	His	Lys	Met	Glu	Met	Lys	Gln	Ala	Ile	Glu	Leu	Val	Glu	Ser	Gly
580					585						590				
Gly	Met	Gly	Lys	Ile	Pro	Ser	Ala	Val	Ser	Thr	Val	Ser	Met	Gln	Asn
595					600					605					
Ile	His	Pro	Lys	Ser	Leu	Pro	Ser	Glu	Arg	Ile	Leu	Pro	Ala	Leu	Ser
610					615					620					
Lys	Asp	Lys	Glu	Glu	Ile	Arg	Lys	Ile	Leu	Arg	Asn	Asn	Leu	Gln	
625					630				635				640		
Lys	Thr	Arg	Gln	Arg	Leu	Arg	Ser	Tyr	Asn	Arg	His	Thr	Leu	Val	Ala
645					650				655						
Asp	Pro	Tyr	Glu	Glu	Ala	Trp	Asn	Gln	Met	Leu	Leu	Arg	Arg	Gln	Lys
660					665					670					
Ala	Arg	Gln	Leu	Glu	Gln	Lys	Ile	Asn	Asn	Tyr	Leu	Thr	Val	Pro	Ala
675					680					685					
His	Lys	Leu	Asp	Ser	Pro	Thr	Met	Ser	Arg	Ala	Arg	Ile	Gly	Ser	Asp
690					695					700					
Pro	Leu	Ala	Tyr	Glu	Pro	Lys	Glu	Asp	Leu	Pro	Val	Ile	Thr	Ile	Asp
705					710				715			720			
Pro	Ala	Ser	Pro	Gln	Ser	Pro	Glu	Ser	Val	Asp	Leu	Val	Asn	Glu	Glu
725					730					735					
Leu	Lys	Gly	Lys	Val	Leu	Gly	Leu	Ser	Arg	Asp	Pro	Ala	Lys	Val	Ala
740					745				750				750		
Glu	Glu	Asp	Glu	Asp	Asp	Gly	Gly	Ile	Met	Met	Arg	Ser	Lys	Glu	
755					760				765						
Thr	Ser	Ser	Pro	Gly	Thr	Asp	Asp	Val	Phe	Thr	Pro	Ala	Pro	Ser	Asp
770					775				780						
Ser	Pro	Ser	Ser	Gln	Arg	Ile	Gln	Arg	Cys	Leu	Ser	Asp	Pro	Gly	Pro
785					790				795			800			
His	Pro	Glu	Pro	Gly	Glu	Gly	Glu	Pro	Phe	Phe	Pro	Lys	Gly	Gln	
					805				810			815			

<210> SEQ ID NO 41

<211> LENGTH: 4126

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 41

tccatccggc	cgtggtccag	cctatccggc	cttccccctc	tcttcaggac	gcagttcccc	60
cgatttcctt	ctctcagect	gaagttctag	atcctttcgg	tttcctcctt	agtccctcca	120
aggtaacttaa	cccccttttg	cagtttggac	tttgagatgt	tgaatctcta	aaaaacgaaa	180
atacttctct	aaaacatgaa	ctttccttag	agattgcccc	agtctcaccc	catctgcagg	240

-continued

actccttgc	acctatgtcc	caggtcttac	acgcacttcg	ctaaggctcg	aggcctccta	300
ggctggggaa	tcagtatgtat	gcttcgggtgg	tccggcgctct	ggggatttca	tccaccccg	360
atcttccccct	ccttgctgggt	ggtggttgcc	ttgggtggac	tgctaccctgt	tctcaggagc	420
cacggccctcc	ageacacagecc	tactgccagc	accatcagag	gttctgaacc	acccegggaa	480
cgctcaattt	gggatgtcac	cacggcacca	tcagagecctc	tccatcgccc	agatgaccac	540
aatttgcacca	acttaatcat	tgaacatggc	ggtaageccat	ctcggaaaggc	cttcccagtc	600
ctggacattt	actaccctca	cgtgcgcaca	ccctttgaga	tctccctctg	gatttcctg	660
gcctgcctca	tgaagatagg	tttccatgtt	atccccacca	tctcaagcat	cgtccggag	720
agctgcctgc	tgatcgtagt	ggggctgctg	gtggggggcc	tgatcaaggg	tgtcgagag	780
acgeccccct	tcctgcaatc	agacgttctc	ttccctttcc	tgctgccacc	catcatcctg	840
gatgcaggct	acttcctgcc	tctgcggcag	ttcacggaga	acctggcac	catcctgatc	900
tttgctgtgg	ttgggcacact	gtggaatgc	ttcttcctgg	gtggcctct	gtacgcctg	960
tgcctgggtgg	gcggcgagca	gatcaacaac	attggcctgc	tggacaccct	gtctttggc	1020
agcatcatct	ccgcgtggaa	ccccgtggct	gtgctggccg	tcttcgagga	gatccacatc	1080
aacgaactgc	tacacatct	tgtcttcggg	gagtccctgc	tcaacgtatgc	cgtcactgt	1140
gtcctctatac	acccctttga	ggaatttgcc	agctatgact	ctgtggcat	ctcgacatc	1200
ttcctcggtt	tcctgagett	cttcgtggtg	gcctctgggt	gggtgtttgt	gggcgtggc	1260
tacggggtca	tcgcggcttt	cacctcccg	tttacctccc	acatccgggt	catcgagccg	1320
ctcttcgtct	tcctctacag	ctacatggct	tacctgtcag	ctgagcttt	ccacctgtcc	1380
ggaatcatgg	ccctcategc	ctcaggagta	gtgatgcgc	cctatgtgga	ggccaacatc	1440
tcccacaata	cccacaccac	catcaagtac	ttccctgaaga	tgtggagcag	cgtcagttag	1500
accctcatct	tcatcttctt	cggcgtctcc	actgtggccg	gctccacca	gtggaaactgg	1560
accttcgtca	tcagcacccct	gtcttctgc	ctcatcgccc	gggtgtctgg	tgtcctggc	1620
ctgacctgggt	tcatcaacaa	gttccgcatt	gtcaagctga	cccccaagga	ccagttcatc	1680
attgcctatg	ggggcctgcg	aggggcacatc	gccttctcgc	tgggctacct	cctggacaaag	1740
aagcacccct	ccatgtgtga	cctgttccctc	accgcacatc	tcaccgtcat	cttttcacc	1800
gtctttgtgc	agggcatgac	cattcggccc	ctgggtcgacc	tgttggccgt	gaagaaaaag	1860
caagaaacaa	agcgctccat	caacgaggag	atccacacac	agttcctgga	ccaccttctg	1920
acaggcatcg	aggacatctg	tggtcattat	ggccaccatc	actggaaagg	caagctcaac	1980
cgttttaata	agaagtatgt	gaagaagtgc	ctgatagcag	gagagcgctc	caaagagccc	2040
cagctcattt	ctttctacca	caagatggag	atgaagcagg	ccattgagct	ggtggagagc	2100
ggaggcatgg	gcaagatccc	atctgctgtc	tccactgtct	ctatgcagaa	catccacccc	2160
aaggctgtga	cttcagaccc	catcttgc	gcactgtcca	aggacaagg	ggaagagatc	2220
cgcaaaatct	tgaggagcaa	tctgcagaaa	accggcaac	ggctgcggtc	ctataacaga	2280
cacacgctgg	tggccgaccc	ctacgaggaa	gcctggacc	agatgctgt	ccggaggcag	2340
aaggccccgc	agctggagca	gaagatcacc	aactacctga	cagtaccggc	ccacaagctg	2400
gactcgccca	ccctgtctcg	ggcccgatca	ggctcagacc	cactggccta	tgaacccaag	2460
gcagacctgc	ctgtcatcac	cattgacccg	gcctccccac	agtccctgta	gtctgtggac	2520

-continued

ctggtaacg aggagttgaa gggcaaggtc ctagggctaa accggggtcc cagggtgact	2580
ccggaggagg aggaggaaga tgaggatggc atcatcatga ttccggagcaa ggagccctca	2640
tccccaggca ccgacgacgt ctccacccct gnatcaagt acagccccag ctcccagaga	2700
ataacaacgt gcctcagtga cccagggccc caccctgagc ctggggaggg agagecttc	2760
atccccaaag gacagtagag ccaggctagt gggtgtccca cagactctcc cacgagagt	2820
ggggctgggc gctcctcttg tccctgacc tggattggct ctgcctccctt cctccgcatg	2880
gcagctgggc ccacagcccc cactgcagca cggcttgccc ctgcctccctg ggaagggtcc	2940
tctctacag aactgccttc ccaatccatc catttggcg atggctgggc tgctgaggca	3000
ggccccgccc catctctaga tccaggetct cccatagtgg aaccagggt cctcaaggcc	3060
ctccttcctt ccagcccttc ctcatccaga ccaactctcc atccttgatc atcattcta	3120
accaaagagc cactggctcc actgtggtcc cagcaggaaa ggaggagct gagccttccg	3180
tggccgtgac tggggcccttg tatcagagga cagcaggcct gtggctggga agcagacccg	3240
tgaccatag tagagctgcc cacggcagcg ccacggttct cccgctagcc tcggtgctct	3300
ggccgcacc accgactgtcc ttactcagag ctgcgggctg agggcgtctc tgagttctc	3360
tgctgggcgc tggggcacat tcaacagaga cgtgattggg gtgatgagga ccactgggcc	3420
ttagtggggc ttacacggga gggactgtcc gaggctctgg aaagtgtctt ctttttacac	3480
ttaggtttatt tgcacctcac tgtaagtatg attctgtttg cacagaggca ctctgtttt	3540
cagacccttt gagtccccca tgccagtcag agctctgccc ttccccacc ttcccttaaa	3600
tggAACCTT gctgctgggg gcccctgccc acagcaccga ttctgatgtg gcccaacctg	3660
tcatgactca agetcaccag cagggggcct cccggtaacct cccttgctc tgggctgtct	3720
tttacttctg ttggccagtt ctaccaccac cccttatctg gaccactctc ttctctgcat	3780
tctgttagagt agaaaagtact ggacccccc agaggcctt gtcacagggc acacactccc	3840
taccggccacc atggctgggg caagccccatg ggtatctgcc caaccctgg gcactccctc	3900
cccaggatcg tgcaataata gtcagagggg ccctccagg ggtggggcc atgggtgcat	3960
ggcccagctg tctcctctcc atgtaagtgc tgttttggc aagagttgcc aggcagggtg	4020
acatcgacaa ccatgtacaa agccaccgca gctgctgccc ctctgccc tgcataaaagg	4080
aaactgaacc ttttcatat tctaataaat caacgtgagt ttttga	4126

<210> SEQ ID NO 42

<211> LENGTH: 820

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 42

Met Met Leu Arg Trp Ser Gly Val Trp Gly Phe His Pro Pro Arg Ile			
1	5	10	15

Phe Pro Ser Leu Leu Val Val Ala Leu Val Gly Leu Leu Pro Val			
20	25	30	

Leu Arg Ser His Gly Leu Gln His Ser Pro Thr Ala Ser Thr Ile Arg			
35	40	45	

Gly Ser Glu Pro Pro Arg Glu Arg Ser Ile Gly Asp Val Thr Thr Ala			
50	55	60	

Pro Ser Glu Pro Leu His Arg Pro Asp Asp His Asn Leu Thr Asn Leu			
65	70	75	80

-continued

Ile	Ile	Glu	His	Gly	Gly	Lys	Pro	Ser	Arg	Lys	Ala	Phe	Pro	Val	Leu
85		90		95											
Asp	Ile	Asp	Tyr	Pro	His	Val	Arg	Thr	Pro	Phe	Glu	Ile	Ser	Leu	Trp
100		105		105		110									
Ile	Leu	Leu	Ala	Cys	Leu	Met	Lys	Ile	Gly	Phe	His	Val	Ile	Pro	Thr
115		120		125											
Ile	Ser	Ser	Ile	Val	Pro	Glu	Ser	Cys	Leu	Leu	Ile	Val	Val	Gly	Leu
130		135		135		140									
Leu	Val	Gly	Gly	Leu	Ile	Lys	Gly	Val	Gly	Glu	Thr	Pro	Pro	Phe	Leu
145		150		150		155		155		160					
Gln	Ser	Asp	Val	Phe	Phe	Leu	Phe	Leu	Leu	Pro	Pro	Ile	Ile	Leu	Asp
165		170		170		175									
Ala	Gly	Tyr	Phe	Leu	Pro	Leu	Arg	Gln	Phe	Thr	Glu	Asn	Leu	Gly	Thr
180		185		185		190									
Ile	Leu	Ile	Phe	Ala	Val	Val	Gly	Thr	Leu	Trp	Asn	Ala	Phe	Phe	Leu
195		200		200		205									
Gly	Gly	Leu	Leu	Tyr	Ala	Val	Cys	Leu	Val	Gly	Gly	Glu	Gln	Ile	Asn
210		215		215		220									
Asn	Ile	Gly	Leu	Leu	Asp	Thr	Leu	Leu	Phe	Gly	Ser	Ile	Ile	Ser	Ala
225		230		230		235		235		240					
Val	Asp	Pro	Val	Ala	Val	Leu	Ala	Val	Phe	Glu	Glu	Ile	His	Ile	Asn
245		250		250		255									
Glu	Leu	Leu	His	Ile	Leu	Val	Phe	Gly	Glu	Ser	Leu	Leu	Asn	Asp	Ala
260		265		265		270									
Val	Thr	Val	Val	Leu	Tyr	His	Leu	Phe	Glu	Glu	Phe	Ala	Ser	Tyr	Asp
275		280		280		285									
Ser	Val	Gly	Ile	Ser	Asp	Ile	Phe	Leu	Gly	Phe	Leu	Ser	Phe	Phe	Val
290		295		295		300									
Val	Ala	Leu	Gly	Gly	Val	Phe	Val	Gly	Val	Val	Tyr	Gly	Val	Ile	Ala
305		310		310		315		315		320					
Ala	Phe	Thr	Ser	Arg	Phe	Thr	Ser	His	Ile	Arg	Val	Ile	Glu	Pro	Leu
325		330		330		335									
Phe	Val	Phe	Leu	Tyr	Ser	Tyr	Met	Ala	Tyr	Leu	Ser	Ala	Glu	Leu	Phe
340		345		345		350									
His	Leu	Ser	Gly	Ile	Met	Ala	Leu	Ile	Ala	Ser	Gly	Val	Val	Met	Arg
355		360		360		365									
Pro	Tyr	Val	Glu	Ala	Asn	Ile	Ser	His	Lys	Ser	His	Thr	Thr	Ile	Lys
370		375		375		380									
Tyr	Phe	Leu	Lys	Met	Trp	Ser	Ser	Val	Ser	Glu	Thr	Leu	Ile	Phe	Ile
385		390		390		395		395		400					
Phe	Leu	Gly	Val	Ser	Thr	Val	Ala	Gly	Ser	His	Gln	Trp	Asn	Trp	Thr
405		410		410		415									
Phe	Val	Ile	Ser	Thr	Leu	Leu	Phe	Cys	Leu	Ile	Ala	Arg	Val	Leu	Gly
420		425		425		430									
Val	Leu	Val	Leu	Thr	Trp	Phe	Ile	Asn	Lys	Phe	Arg	Ile	Val	Lys	Leu
435		440		440		445									
Thr	Pro	Lys	Asp	Gln	Phe	Ile	Ile	Ala	Tyr	Gly	Gly	Leu	Arg	Gly	Ala
450		455		455		460									
Ile	Ala	Phe	Ser	Leu	Gly	Tyr	Leu	Leu	Asp	Lys	Lys	His	Phe	Pro	Met
465		470		470		475		475		480					

-continued

```

Cys Asp Leu Phe Leu Thr Ala Ile Ile Thr Val Ile Phe Phe Thr Val
485          490          495

Phe Val Gln Gly Met Thr Ile Arg Pro Leu Val Asp Leu Leu Ala Val
500          505          510

Lys Lys Lys Gln Glu Thr Lys Arg Ser Ile Asn Glu Glu Ile His Thr
515          520          525

Gln Phe Leu Asp His Leu Leu Thr Gly Ile Glu Asp Ile Cys Gly His
530          535          540

Tyr Gly His His His Trp Lys Asp Lys Leu Asn Arg Phe Asn Lys Lys
545          550          555          560

Tyr Val Lys Lys Cys Leu Ile Ala Gly Glu Arg Ser Lys Glu Pro Gln
565          570          575

Leu Ile Ala Phe Tyr His Lys Met Glu Met Lys Gln Ala Ile Glu Leu
580          585          590

Val Glu Ser Gly Gly Met Gly Lys Ile Pro Ser Ala Val Ser Thr Val
595          600          605

Ser Met Gln Asn Ile His Pro Lys Ala Val Thr Ser Asp Arg Ile Leu
610          615          620

Pro Ala Leu Ser Lys Asp Lys Glu Glu Ile Arg Lys Ile Leu Arg
625          630          635          640

Ser Asn Leu Gln Lys Thr Arg Gln Arg Leu Arg Ser Tyr Asn Arg His
645          650          655

Thr Leu Val Ala Asp Pro Tyr Glu Glu Ala Trp Asn Gln Met Leu Leu
660          665          670

Arg Arg Gln Lys Ala Arg Gln Leu Glu Gln Lys Ile Thr Asn Tyr Leu
675          680          685

Thr Val Pro Ala His Lys Leu Asp Ser Pro Thr Leu Ser Arg Ala Arg
690          695          700

Ile Gly Ser Asp Pro Leu Ala Tyr Glu Pro Lys Ala Asp Leu Pro Val
705          710          715          720

Ile Thr Ile Asp Pro Ala Ser Pro Gln Ser Pro Glu Ser Val Asp Leu
725          730          735

Val Asn Glu Glu Leu Lys Gly Lys Val Leu Gly Leu Asn Arg Gly Pro
740          745          750

Arg Val Thr Pro Glu Glu Glu Asp Glu Asp Gly Ile Ile Met
755          760          765

Ile Arg Ser Lys Glu Pro Ser Ser Pro Gly Thr Asp Asp Val Phe Thr
770          775          780

Pro Gly Ser Ser Asp Ser Pro Ser Ser Gln Arg Ile Gln Arg Cys Leu
785          790          795          800

Ser Asp Pro Gly Pro His Pro Glu Pro Gly Glu Gly Glu Pro Phe Ile
805          810          815

Pro Lys Gly Gln
820

```

<210> SEQ_ID NO 43

<211> LENGTH: 5069

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 43

aagttcctg aaacggagaa acataaacag gcattgtgg gatcacccat acatcactct 60

-continued

gttttgctg acctttccg gtaatttcaa aacaaacccg gtctcgaagc ggagatccgg	120
cgataattac cgcagaaaata aaccataca cgagacgtag aaccagccgc acatggccgg	180
agaaaactct gcgagaattt cgtaaactcg cgcgcatgc atctgttattt cctaatgcgg	240
cacttccagg cctcgagacc tctgacatgc ttttgacagg aatagacatt ttcaaatgt	300
tatccatatg ccttcgggt tttttccctt cctttccat catgaaaaat ctctcgagac	360
cgtttatcca ttgtttttt ttgttctttt ccctctgttc acagaaaatc tgaagaagct	420
atagtagaac tatgagctt ttttgttctt ttgtttttt tacctctgtg	480
gaaattgtta ctctcacact cttagtgc ttgtttgtt ttgtttattt caattatgac	540
cggtgacgaa acgtggtega tggtgggtac cgcttatgct cccctccatt agttcgatt	600
atataaaaaag gccaaatattt gtattttttt caaatgtcct atcattatcg tctaacatct	660
aatttctctt aaatttttttcc ttgttcttc ctataacacc aatagtggaa atctttttt	720
cttctatatc tacaaaaaact ttgtttttctt atcaacctcg ttgataaattt ttgtttttttaa	780
caatcgtaa taattaatta attggaaaat aaccattttt tctctttttt atacacacat	840
tcaaaaagaaa gaaaaaaaaat ataccccagc tagttaaaga aaatcattga aaagaataag	900
aagataagaa agatttaaattt atcaaacaat atcaatatcg ctgatacatac atcctttca	960
tcatcctctt cagcatcttc tggttcagct catcagccaa ctcaagaaaa gcctgtcaag	1020
acttacgatg acgctgcattc tgaatcttctt gacgatgacg atatcgatgc ttatcgaa	1080
gaactacaat ctaatcacgg tgtcgacgac gaagacagtg ataacgatgg tccagttgcc	1140
gccgggtgaag cttagaccgt tccagaagaa tatttacaaa ctgaccatc ttacgggtta	1200
acttccgatg aagttttgaa aagaagaaaag aagtacgggtt tgaatcaaat ggctgacgaa	1260
aaagaatctt tagtcgttaa gttcgtttagt ttgttctgtc gtccaaatca attcgtttagt	1320
gaagccgctg ctatttggc tgccgggttgc tccgatttggg tcgatttgc tggttatctgt	1380
ggtttgcattaa tgtaaacgc tggtgttgggtt tcggttcaag aattccaacgc tggttctatt	1440
gtcgatgaat tgaaaaagac ttggcttaac actgctgttg ttatcagaga cggtaatttgc	1500
gttgaattc cagccaaacga agtcgtcccc ggtgatattt tgcaatttggg agatggact	1560
gttatccaa ctgatggctg tattgtcaactt gaagactgtt tcttgcaaat cgatcaatct	1620
gctattactg gtatctttt ggctgtcgac aaacatttacg gtgaccaaac ttcttttct	1680
tccactgtta agagaggtga aggtttcatg gttgttaccg ctaccggta caacacttcc	1740
gttggtagag ctgctgtttt ggtaacaaa gcccgtgggtt gtcaagggtca ttctactgaa	1800
gttttgaacg gtattggat tatcttattt gttttggtaa ttgccactttt gttgttggtc	1860
tggactgctt gtttctacag aaccaacgggtt attgttagaa tctttagata cactcttaggt	1920
attactatta ttgggtgtccc agtcgggtttt ccagctgtcg ttaccaccac tatggcgtc	1980
ggtgctgtttt acttggctaa gaaacaagcc attgttcaaa agttgtctgc tattgaatcc	2040
ttggctgtgtg tcgaaatctt gtgttccgac aaaaccggta ctgttgcaccaaa gaacaagtttgc	2100
tccttgcacg aaccatacacat tggtaaggtt gtttctccag acgacttgcgtt gttgactgt	2160
tgtttggctg cttccagaaa gaagaaggggtt ttggatgcta ttgataaggc ttcttgaag	2220
tctttgaagc aatatccaaa ggctaaaggac gctttgaccaaa agtacaaggtt ttggaaattc	2280
catccatttgc accctgtctc caagaaggtt actgcccgttgc tcgaatctcc agaagggtgaa	2340

-continued

agaatttgtt tggttaaggg tgctccatta ttcgtcttga agactgtcga agaagatcac 2400
ccaatcccag aagatgtcca tggaaaactac gaaaacaagg ttgctgaaatt ggcttctaga 2460
gggttccgtg ctttaggtgt tgctagaaag agaggtgaag gtcactgggaaatctgggt 2520
gttatgccat gtatggatcc accaagagac gatactgctc aaactgttag cgaaagctaga 2580
caactgggtc taagagttaa gatgttaact ggtgatgctg ttggatttgc taagggaaact 2640
ttagacaat tgggttggg taccaacatt tacaacgctg aaagattagg tcttagtgcc 2700
ggtggtgaca tgccaggttc tgaatttagct gatTTTgtt gaaatgccga tggttcgct 2760
gaagtttcc cacaacataa atacagagtt gttgaaatct tgcaaaacag aggttacttg 2820
gttgctatga ctggtgatgg tggtaacgat gctccatctt tgaagaaggc tgatactgg 2880
attgtctcg aagggtctac tggatgctgc agatctgctg ctgatattgt ttcttggtc 2940
cctggcttat ctgctattat tggatgcttg aagacctcca gacaaatttt tcacagaatg 3000
tactcttacg ttgtttaccg tattgcttt tctctacatt tggaaatctt cttgggtcta 3060
tggattgcta ttttggataa ctcttggac attgatttga ttgtttcat cgctattttc 3120
gtgtgatgtt ctactttggc tattgcttac gataatgctc cttaactctcc aaaggccgtt 3180
aaatggAACc taccaagatt atggggatgt tctattttt tggcatagt tttagctata 3240
ggttcttggaa ttaccttgac tactatgttcc ttaccaaagg gtggattttt cc当地acttc 3300
ggtgctatga acggattttt gttcttgcaat tttccttga ctgaaaactg gttgattttc 3360
attaccagag ctgctggtcc attctggct tctatccat cctggcaatt ggctgggtgcc 3420
gtcttcgtg tcgacatcat cgctaccatg tttacctt tgggttggggtt gtcgtttttttc 3480
tggactgata ttgttactgt cgtccgttgc tggatctggt ctatcggtat cttctgtt 3540
ttgggtgggtt tctactacga aatgtccact tctgaaggct ttgacagatt gatgaacgg 3600
aaggccatga aggaaaagaa gtctaccaga agtgtcgaag acttcatggc tgctatgca 3660
agagtctctca ctcaacacga aaaggaaacc taatcctgtt gaagtagcat ttaatcataa 3720
tttttgcac attttaatca acttgatttt tctggtttaa tttttcttaat ttaatttta 3780
atttttttat caatgggaac tgatacacta aaaagaattt ggagccaaaca agaataagcc 3840
gtttaattcc tactagagtt attttatcca cacacacacc ttaaaaatttt tagatttaat 3900
ggcatcaact cttagotcca cacacacaca cacaccgaag ctgggttgggat tatttgcatt 3960
gatataattt gtttctctgg atggactttt ttctttttt ttctttcttgc gttattttcc 4020
attttaaaat atgaaacgca cacaagtcata aattattctta atagagcaca attcacaaca 4080
cgccacatttc aactttaataa tttttttaga aacactttt ttagtcttaat tcttaatttt 4140
taatataat aatgcacaca cactaatttta ttcatttaattt ttttatttgag taggatttga 4200
aaatatttgg tatcttgca agatgtttgtt atagaggcaca aaagaatcgt cttaatttt 4260
gtcaaggctt tacgtcataa tagttctgc ccagcttcc tataatactt taaagatctc 4320
ttctcggtt ctccattttgg aagtctcgat tacgtttatg cgccccataca gacactcaag 4380
atacacactt acatgaacgtt atacaattt actaacacta cttgaaaata tgaaccacag 4440
tacatcatat taagacgtat tattcgatga ttgaaggccg cctcccgcaat tacctttatc 4500
gatTTTggccg gttatcgca tggaaatttc ttctgtcaca gaaagcaaac aaatcgccag 4560
gccttctac cagtttccctt ttctttagaa gatgtaaaag ctactaaccg cacattactc 4620

-continued

tagatgactc agtttagtct gacccatctat actataacta ccctggcgct atgatgtat	4680
gcgggttcttt tattgcggaa acgaaaattc cgggaccggc gaaatttgcc cggtttgtc	4740
gtaaccggct tcatacgatcg gcttcataatag tagttgaata ctatattaaa cagcagaact	4800
taactcaactc atcacgctgt ttccgctgaa ttttctcaaa atatctaagc agtcaacaaa	4860
tataaagaat attgaaattt gacagtttt tgtcgctatc gattttattt atttgctgtt	4920
ttaaatcatg gtttacactc catccaaggg tccaagaact cttaacgata aggttttga	4980
tgcacatgtt gtccatcaag atgaaaatgg ttcccttttgg ttgtatatcg acagacactt	5040
ggttcatgaa gtcacatctc cacaagtt	5069

<210> SEQ ID NO 44

<211> LENGTH: 918

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 44

Met Thr Asp Thr Ser Ser Ser Ser Ser Ser Ala Ser Ser Val	
1 5 10 15	

Ser Ala His Gln Pro Thr Gln Glu Lys Pro Ala Lys Thr Tyr Asp Asp	
20 25 30	

Ala Ala Ser Glu Ser Ser Asp Asp Asp Ile Asp Ala Leu Ile Glu	
35 40 45	

Glu Leu Gln Ser Asn His Gly Val Asp Asp Glu Asp Ser Asp Asn Asp	
50 55 60	

Gly Pro Val Ala Ala Gly Glu Ala Arg Pro Val Pro Glu Glu Tyr Leu	
65 70 75 80	

Gln Thr Asp Pro Ser Tyr Gly Leu Thr Ser Asp Glu Val Leu Lys Arg	
85 90 95	

Arg Lys Lys Tyr Gly Leu Asn Gln Met Ala Asp Glu Lys Glu Ser Leu	
100 105 110	

Val Val Lys Phe Val Met Phe Phe Val Gly Pro Ile Gln Phe Val Met	
115 120 125	

Glu Ala Ala Ala Ile Leu Ala Ala Gly Leu Ser Asp Trp Val Asp Phe	
130 135 140	

Gly Val Ile Cys Gly Leu Leu Met Leu Asn Ala Gly Val Gly Phe Val	
145 150 155 160	

Gln Glu Phe Gln Ala Gly Ser Ile Val Asp Glu Leu Lys Lys Thr Leu	
165 170 175	

Ala Asn Thr Ala Val Val Ile Arg Asp Gly Gln Leu Val Glu Ile Pro	
180 185 190	

Ala Asn Glu Val Val Pro Gly Asp Ile Leu Gln Leu Glu Asp Gly Thr	
195 200 205	

Val Ile Pro Thr Asp Gly Arg Ile Val Thr Glu Asp Cys Phe Leu Gln	
210 215 220	

Ile Asp Gln Ser Ala Ile Thr Gly Glu Ser Leu Ala Val Asp Lys His	
225 230 235 240	

Tyr Gly Asp Gln Thr Phe Ser Ser Ser Thr Val Lys Arg Gly Glu Gly	
245 250 255	

Phe Met Val Val Thr Ala Thr Gly Asp Asn Thr Phe Val Gly Arg Ala	
260 265 270	

Ala Ala Leu Val Asn Lys Ala Ala Gly Gly Gln Gly His Phe Thr Glu	
---	--

-continued

275	280	285
Val Leu Asn Gly Ile Gly Ile Ile Leu Leu Val	290	295 300
Leu Val Val Trp Thr Ala Cys Phe Tyr Arg Thr Asn Gly Ile Val	305	310 315 320
Arg Ile Leu Arg Tyr Thr Leu Gly Ile Thr Ile Ile Gly Val Pro Val	325	330 335
Gly Leu Pro Ala Val Val Thr Thr Thr Met Ala Val Gly Ala Ala Tyr	340	345 350
Leu Ala Lys Lys Gln Ala Ile Val Gln Lys Leu Ser Ala Ile Glu Ser	355	360 365
Leu Ala Gly Val Glu Ile Leu Cys Ser Asp Lys Thr Gly Thr Leu Thr	370	375 380
Lys Asn Lys Leu Ser Leu His Glu Pro Tyr Thr Val Glu Gly Val Ser	385	390 395 400
Pro Asp Asp Leu Met Leu Thr Ala Cys Leu Ala Ala Ser Arg Lys Lys	405	410 415
Lys Gly Leu Asp Ala Ile Asp Lys Ala Phe Leu Lys Ser Leu Lys Gln	420	425 430
Tyr Pro Lys Ala Lys Asp Ala Leu Thr Lys Tyr Lys Val Leu Glu Phe	435	440 445
His Pro Phe Asp Pro Val Ser Lys Lys Val Thr Ala Val Val Glu Ser	450	455 460
Pro Glu Gly Glu Arg Ile Val Cys Val Lys Gly Ala Pro Leu Phe Val	465	470 475 480
Leu Lys Thr Val Glu Glu Asp His Pro Ile Pro Glu Asp Val His Glu	485	490 495
Asn Tyr Glu Asn Lys Val Ala Glu Leu Ala Ser Arg Gly Phe Arg Ala	500	505 510
Leu Gly Val Ala Arg Lys Arg Gly Glu Gly His Trp Glu Ile Leu Gly	515	520 525
Val Met Pro Cys Met Asp Pro Pro Arg Asp Asp Thr Ala Gln Thr Val	530	535 540
Ser Glu Ala Arg His Leu Gly Leu Arg Val Lys Met Leu Thr Gly Asp	545	550 555 560
Ala Val Gly Ile Ala Lys Glu Thr Cys Arg Gln Leu Gly Leu Gly Thr	565	570 575
Asn Ile Tyr Asn Ala Glu Arg Leu Gly Leu Gly Gly Asp Met	580	585 590
Pro Gly Ser Glu Leu Ala Asp Phe Val Glu Asn Ala Asp Gly Phe Ala	595	600 605
Glu Val Phe Pro Gln His Lys Tyr Arg Val Val Glu Ile Leu Gln Asn	610	615 620
Arg Gly Tyr Leu Val Ala Met Thr Gly Asp Gly Val Asn Asp Ala Pro	625	630 635 640
Ser Leu Lys Lys Ala Asp Thr Gly Ile Ala Val Glu Gly Ala Thr Asp	645	650 655
Ala Ala Arg Ser Ala Ala Asp Ile Val Phe Leu Ala Pro Gly Leu Ser	660	665 670
Ala Ile Ile Asp Ala Leu Lys Thr Ser Arg Gln Ile Phe His Arg Met	675	680 685

-continued

Tyr Ser Tyr Val Val Tyr Arg Ile Ala Leu Ser Leu His Leu Glu Ile
690 695 700

Phe Leu Gly Leu Trp Ile Ala Ile Leu Asp Asn Ser Leu Asp Ile Asp
705 710 715 720

Leu Ile Val Phe Ile Ala Ile Phe Ala Asp Val Ala Thr Leu Ala Ile
725 730 735

Ala Tyr Asp Asn Ala Pro Tyr Ser Pro Lys Pro Val Lys Trp Asn Leu
740 745 750

Pro Arg Leu Trp Gly Met Ser Ile Ile Leu Gly Ile Val Leu Ala Ile
755 760 765

Gly Ser Trp Ile Thr Leu Thr Met Phe Leu Pro Lys Gly Gly Ile
770 775 780

Ile Gln Asn Phe Gly Ala Met Asn Gly Ile Met Phe Leu Gln Ile Ser
785 790 795 800

Leu Thr Glu Asn Trp Leu Ile Phe Ile Thr Arg Ala Ala Gly Pro Phe
805 810 815

Trp Ser Ser Ile Pro Ser Trp Gln Leu Ala Gly Ala Val Phe Ala Val
820 825 830

Asp Ile Ile Ala Thr Met Phe Thr Leu Phe Gly Trp Trp Ser Glu Asn
835 840 845

Trp Thr Asp Ile Val Thr Val Val Arg Val Trp Ile Trp Ser Ile Gly
850 855 860

Ile Phe Cys Val Leu Gly Gly Phe Tyr Tyr Glu Met Ser Thr Ser Glu
865 870 875 880

Ala Phe Asp Arg Leu Met Asn Gly Lys Pro Met Lys Glu Lys Lys Ser
885 890 895

Thr Arg Ser Val Glu Asp Phe Met Ala Ala Met Gln Arg Val Ser Thr
900 905 910

Gln His Glu Lys Glu Thr
915

<210> SEQ ID NO 45
<211> LENGTH: 8526
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

agacggcgcc ggccccgt aatgcgggg atcgctcccc cggggccgct gagcctgcgc 60
ccagtgcggc gagccccggc cccggccgg tccgcgcggaa gcaggcggccccc 120
ggccggccgg ggggaccggc agttccccca caggcaacgt gaggagagcc tggccggaga 180
agcaggatga gaagatggca aacttcctat tacctcgggg caccaggcggc ttccggcggt 240
tcacacggga gtccctggca gccatcgaga agcgcatggc ggagaagcaa gccccgggt 300
caaccacctt gcaggagggc cgagggggc tgcccgaggaa ggaggctccc cggccccc 360
tggacctgca ggccctccaaa aagctgccag atctctatgg caatccaccc caagagctca 420
tcggagagcc cttggaggac ctggaccctt tctatagcac ccaaaagact ttcatcgta 480
tgaataaagg caagaccatc ttccgggttca gtgccaccaa cgccttgat gtcctcagtc 540
ccttccaccc catccggaga ggggctgtga agattcttgtt tcactcgctc ttcaacatgc 600
tcatcatgtg caccatcctc accaactgcg tggtcatggc ccagcacgac cctccaccc 660

-continued

ggaccaagta tgcgagttac accttcaccg ccatttacac ctttgagtct ctggtaaga	720
ttctggctcg aggcttctgc ctgcacgcgt tcactttcct tcgggaccca tggaactggc	780
tggactttatcg tggattatac atggcataaca caactgaatt tgggacactg ggcaatgtct	840
caccccttaacg cacccctcga gtcctccggg ccctgaaaac tatatcagtc atttcagggc	900
tgaagaccat cgtggggggcc ctgatccagt ctgtgaagaa gctggctgat gtgatggtcc	960
tcacagtctt ctgcctcagc gtctttgcc tcatcgccct gcagcttcc atggcaacc	1020
taaggcacaa gtgcgtgegc aacttcacag cgctcaacgg caccaacggc tccgtggagg	1080
ccgacggctt ggtctggaa tccctggacc ttacccatcg tcatccgaa aattacctgc	1140
tcaagaacgg cacccctgtat gtgttactgt gtggaaacag ctctgacgct gggacatgtc	1200
cgaggaggcta ccgggtgcata aaggcaggcg agaaccggc ccacggctac accagttcg	1260
attcccttgc ctggggcttt ctgcactct tcggctgtat gacgcaggac tgctggagc	1320
gcctctatca gcagaccctc aggtcccgag ggaagatcta catgatttc ttcatgttt	1380
tcatcttccct ggggtccctc tacctggtga acctgatcct ggccgtggc gcaatggcct	1440
atgaggagca aaaccaagcc accatcgctg agaccgagga gaaggaaaag cgcttcagg	1500
aggccatgga aatgctcaag aaagaacacg aggcctcac catcagggt gtggataccg	1560
tgtcccgttag ctccctggag atgtccccctt tggcccccagt aaacaggcat gagagaagaa	1620
gcaagaggag aaaacggatg tttcaggaa ctgaggagtg tggggaggac aggctcccc	1680
agtctgactc agaagatggt cccagagcaa tgaatcatct cagcctcacc cgtggctca	1740
gcaggacttc tatgaagcca ctttccagcc gcggggagcat tttcaccttt cgcaaggcgag	1800
acctgggttc tgaagcagat tttcagatg atgaaaacacg cacagggggg gagagcgaga	1860
gccaccacac atcaactgtgt gtgcctggc ccctgcgcgg gaccagtgcc caggacacgc	1920
ccagtcccgg aacctcggtt cctggccacg ccctccatgg caaaaagaac agcactgtgg	1980
actgcaatgg ggtggctctca ttactggggg caggcgaccc agaggccaca tccccaggaa	2040
gccacccctt ccgcctgtgt atgctagacg acccggcaga cacgaccacg ccatcgagg	2100
agccaggcgcc gccccagatg ctgacccccc aggctccgtg tggatggc ttcgaggagc	2160
caggagcacg gcagcggggcc ctcagcgcag tcagcgtcct caccagcgca ctggaaaggt	2220
tagaggagtc tcgcacacaag tggccacat gctggaaaccc tctcgccctcg cgctacatgt	2280
tctggggatgt ctgcctcgctg tggatgtcca tcaagcaggg agtgaagttt gtggatgttt	2340
acccgtttac tgcactcacc atcaactatgtt gcacatgtact caacacactc ttcatggcgc	2400
tggagcaacta caacatgaca agtgaattcg aggagatgtt gcagggtcgaa aacctggct	2460
tcacagggtt ttcacacgca gagatgttca tcaagatcat tgcctcgac ccctactact	2520
acccatcaaca ggggtggaaat atccatcgaca gcatcatcgat catccttgcgt ctcatggagc	2580
tggggctgtc ccgcgtgcac aacctgtcggt tgctgcgcgc cttccgcgtt ctgcgggtct	2640
tcaagctggc caaatcatgg cccaccctga acacactcat caagatcatc gggaaactcag	2700
tggggggactt ggggaacccgtt acactgggtgc tagccatcat cgtgttccatc ttgtgttgg	2760
tggggcatgca gctctttggc aagaactactt cggagctgag ggacagcgac tcaggctgc	2820
tgcctcgctg gcacatgtat gacttcttc atgccttcat ctcatcttc cgcacatct	2880
gtggagagtg gatcgagacc atgtggactt gcatggaggt gtcggggcag tcattatgcc	2940

-continued

tgcgttgtt	cttgcttgtt	atggtcattt	gcaacccttgtt	ggtccctgaat	cttttcttgg	3000
ccttgctgtt	cagtccttc	agtgcagaca	acctcacagc	ccctgatgag	gacagagaga	3060
tgaacaacct	ccagactggcc	ctggcccgca	tccagagggg	cctgcgcctt	gtcaagcgga	3120
ccacacctgg	tttctgctgt	ggtctcttc	ggcagcgccc	tcagaagccc	gcageccctt	3180
cgccccagg	ccagactgccc	agctgcattt	ccacccctta	ctccccgcca	ccccccagaga	3240
cggagaagg	gcctccacc	cgcaaggaaa	cacgggttga	ggaaggcgag	caaccaggcc	3300
agggcacc	cggggatcca	gagcccgtgt	gtgtgccat	cgctgtggcc	gagttagaca	3360
cagatgacca	agaagaagat	gaggagaaca	gcctggcac	ggaggaggag	tccagcaagc	3420
aggaatccca	gcctgtgtcc	ggtgcccaag	aggccctcc	ggattccagg	acctggagcc	3480
aggtgtcago	gaetgcctcc	tctgaggcc	aggccagtgc	atctcaggcc	gactggcgcc	3540
agcagtggaa	agcggaaaccc	caggccccag	ggtgccgtga	gaccccgag	gacagtgtct	3600
ccgaggccag	cacagcagac	atgaccaaca	ccgctgagct	cctggagcc	atccctgacc	3660
tccggccagg	tgtcaaggac	ccagaggact	gtttcaactga	aggctgtgtc	cgcgctgtc	3720
cctgctgtgc	ggtgacacc	acacaggccc	cagggaaagg	ctgggtggcg	ttgcgcaaga	3780
cctgctacca	catcggtgg	cacagctgg	tgcagacatt	catcatctt	atgatctac	3840
tcagcagtgg	agcgctggcc	tgcaggaca	tctaccta	ggagcggaa	accatcaagg	3900
ttctgcttga	gtatgcccac	aagatgttca	catatgtctt	cgtgctggag	atgctgctca	3960
agtgggtggc	ctacggcttc	aagaagtact	tcaccaatgc	ctgggtgtgg	ctcgacttcc	4020
tcatcgtaga	cgtctctcg	gtcagccctgg	tggccaacac	cctggcttt	gccgagatgg	4080
gccccatcaa	gtcactgccc	acgctgcgtg	cactccgtcc	tctgagact	ctgtcacgg	4140
ttgagggcat	gagggtggtg	gtcaatgccc	tggtggggcc	catccgtcc	atcatgaacg	4200
tcctccctgt	ctgcctcata	ttctggctca	tcttcagcat	catggcggt	aaccttttgc	4260
cggggaaagtt	tgggaggtgc	atcaaccaga	cagaggaga	cttgccttttgc	aactacacca	4320
tcgtgaacaa	caagagccag	tgtgagtcct	tgaacttgc	cggagaatttgc	tactggacca	4380
aggtgaaagt	caactttgc	aacgtgggggg	ccgggtacct	ggcccttcttgc	cagggtggccaa	4440
catttaaagg	ctggatggac	attatgtatg	cagctgtgga	ctccagggggg	tatgaagagc	4500
agcctcagtg	ggaataacaac	ctctacatgt	acatctatttgc	tgtcattttc	atcattttgc	4560
ggtctttctt	caccctgaac	ctctttatgt	gtgtcatcat	tgacaacttcc	aaccaacaga	4620
agaaaaaaagtt	agggggccag	gacatcttca	tgacagagga	gcagaagaag	tactacaatgt	4680
ccatgaagaa	gctgggcctcc	aagaagcccc	agaagccat	cccacggccc	ctgaacaagt	4740
accaggccctt	catattcgac	attgtgacca	agcaggccctt	tgacgtcacc	atcatgtttc	4800
tgtatgtctt	gaatatggtg	accatgtatgg	tggagacaga	tgaccaaagt	cctgagaaaa	4860
tcaacatctt	ggccaagatc	aacctgtctt	ttgtggccat	cttcacacggc	gagtgtatttgc	4920
tcaagctggc	tgcctgcgc	cactactact	tcaccaacag	ctggaatatac	tgcgacttgc	4980
tgggtgtcat	cctctccatc	gtgggcacttgc	tgctctcgga	catcatccag	aagtacttcttgc	5040
tctccccgac	gctcttccga	gtcatccgc	tggcccgaaat	aggccgcac	ctcagactga	5100
tccgaggggc	caaggggatc	cgcacgcgtgc	tctttgcctt	catgtatgtcc	ctgcctgc	5160
tcttcaacat	cgggcgtgtc	cttttctctcg	tcatgttcat	ctactccatc	tttggcatgg	5220

-continued

ccaaacttcgc ttatgtcaag tgggaggctg gcatcgacga catgttcaac ttccagacct	5280
tcgccaacag catgctgtgc ctcttccaga tcaccacgtc ggccggctgg gatggctcc	5340
tcaaaaaaaaaaaaaacttccat cctcaacact gggccgcctt actgcgaccc cactctgccc aacagaatcg	5400
gctctcgaaaaa gggactgcggg agccccggcc tgggcatctt cttcttacc acctacatca	5460
tcatctccctt cctcatcgta gtcaacatgt acattgccat catcctggag aacttcagcg	5520
tggccacggaa ggagagcacc gagccccctga gtgaggacga cttcgatatg ttcttatgaga	5580
tctggagaaa atttgaccca gaggccactc agtttattga gtattcggtc ctgtctgact	5640
ttggccatgc cctgtctgag ccactccgtt tggccaagcc caaccagata agcctcatca	5700
acatggaccc gccccatggta agtggggacc gcatccatgt catggacatt ctcttcgcct	5760
tcaccaaaaag ggtctctgggg gagtctctgggg agatggacgc cctgaagatc cagatggagg	5820
agaagttcat ggcagccaac ccatccaaga ttccttacga gcccatttacc accacactcc	5880
ggcgcaagca cgaagagggtg tggccatgg ttatccagag agcctccgc aggcacactgc	5940
tgcaacgcctc tttgaagcat gcttccttcc ttttcgtca gcaggcgcc agcggccct	6000
ccgaagagga tggcccttggag cgagagggcc tcatacgatc cgtgtatggat gagaacttct	6060
cccgacccctt tggcccaaccc tccagcttcc tcatacttcc cacttccttc ccacccttcc	6120
atgacagtgtt cactagagcc accagcgata acctccaggt ggggggtct gactacagcc	6180
acagtgaaga ttcggccgac ttccccctt ctcggacag ggaccgtgag tccatcggt	6240
gagcctcgcc ctggctggcc aggacacact gaaaagcagc ctttttaccat atggcaaacc	6300
taaatgcagt cagtcacaaa ccagccttggg gccttcctgg ctggggggat aagaatggg	6360
cctcagcccccc gggatcaac caggcagatg tctgtggccgc cgcgtggaca gccggagcag	6420
ttggcctgtt cttggaggcc tcagatagac ctgtgacccgt gtctggtcag gcaatgcct	6480
gcggctctgg aaagcaactt catcccagct gctgaggcga aatataaaac tgagactgta	6540
tatgttgcata atggggcttc ataaatttat tatattttat atttttttat ttgagcaaaag	6600
aactaaggat tttccatgg acatggccgac caattcacgc tgcgttcttca taaccctgaa	6660
caagagtgtc tatggagccgac ccggaaatgtct gttctcaaag cagaagtggaa atccagtgt	6720
gctcccacag gtcttactg cccagggtc gaatggggtc cccctccac ttgacccgt	6780
atgctggggag ggctgaaccc ccactcacac aagcacacac acacacttcc cacacacgg	6840
ggccagacac agggcggtgg acccaggctc ccagcctaag ggagacagggc ctttcctgc	6900
cgccccccca aggtggggat tcttgcaccc gggggctact ctggcccccatttgcata	6960
aggccccattt ttccccctgt gtttcacgc aggtcatatt gtcagtcata caaaaataaa	7020
aggcttccatcaggaggatgg cctgggtccc agggctggcc ctggactgatgatgttgcct	7080
tttcttcccccc tcttgcataatgtttaacaa aacccaaagga cacaagggtg caagccccat	7140
tcacggccctg gcatgcacgt tgcgttgccttgccttgcataatgttgccttgcataatgttgcct	7200
ccatcggaag agagggtgtga gccatgggggg tttggggctta agaagttcac cagccctgag	7260
ccatggccgc ccctcagectt gcttgcacatcggaggaaactgg cgtatccca gggctctctg	7320
gaccatacgc ggaggaggtt tctgtgtggt ctccagctcc tctccagacaca cagagacatg	7380
ggagtggggaa gggggatgg gccctgcgccttgcgtgcaggaaaggatgg tcaggcccg	7440
ttctcgatgc ctttagagggg aatgaaccat ggcaccccttgcgtgcaggaaaggatgg tcaggcccg	7500

-continued

aggcccagcc tctctggctc agccccggat cctgatggca cccacacaga ggacctctt	7560
ggggcaagat ccaggtggc ccataaggct ttttttcagg gaaaaatatt	7620
ttactatgtcc aatcaccccc aggaccttt cagctgtga caatcttatt tagcatatgc	7680
aaatctttta acatagagaa ctgtcacccctt gaggtAACAG ggtcaactgg cgaaggctga	7740
gcaggcaggg gcttggctgc cccattccag ctctccatg gagccccctcc accggcgca	7800
tgcctcccaag gcccacccatg tctcacctgc cggtctggg ctggctgtc ctaacctacc	7860
tgcggcggact gtccggagggc tggacatttg tggcagtgtc gaagggggca ttggccggcga	7920
gtaaaagtatt atgtttcttc ttgtcaccc agttcccttg gtggcaaccc cagacccaaac	7980
ccatgccccct gacagatcta gttctttctt cctgtgttcc cttttagtcc agtgtgggac	8040
acgggttaaac tggccaccccg acatttctcc aagtggaaat cctatttttg tagatctcca	8100
tgcctttgtc tcaaggcttg gagaggtatg tgcccccttgggtgtgtcac cgccgtctac	8160
acaggcaggaa atgcgggtgg gaggcaggcc gggctgcaccccgatggc cggaggaga	8220
ctgtgggtttt tgggtgtgtg gacagccccgg gagcttttag acaggtgcct ggggtggct	8280
gcagacgggtg tgggtggggg tggggaggta gctagacccaa acccttagct tttagcctgg	8340
ctgtcacccct tttaatttcc agaactgcac aatgaccaggc aggaggaaag gacagacatc	8400
aagtggccaga tgggtgtctga actaatcgag cacttctcac caaactcat gtataataaa	8460
aatacatatt tttaaaacaa accaataaat ggcttacatg aaaaaaaaaa aaaaaaaaaa	8520
aaaaaaaaa	8526

<210> SEQ ID NO 46
<211> LENGTH: 2015
<212> TYPE: PRT
<213> ORGANISM: *Homo sapiens*

<400> SEQUENCE: 46

Met	Ala	Asn	Phe	Leu	Leu	Pro	Arg	Gly	Thr	Ser	Ser	Phe	Arg	Arg	Phe
1				5					10				15		
Thr	Arg	Glu	Ser	Leu	Ala	Ala	Ile	Glu	Lys	Arg	Met	Ala	Glu	Lys	Gln
20					25					30					
Ala	Arg	Gly	Ser	Thr	Thr	Leu	Gln	Glu	Ser	Arg	Glu	Gly	Leu	Pro	Glut
35						40				45					
Glu	Glu	Ala	Pro	Arg	Pro	Gln	Leu	Asp	Leu	Gln	Ala	Ser	Lys	Lys	Leu
50					55					60					
Pro	Asp	Leu	Tyr	Gly	Asn	Pro	Pro	Gln	Glu	Leu	Ile	Gly	Glu	Pro	Leu
65						70				75				80	
Glu	Asp	Leu	Asp	Pro	Phe	Tyr	Ser	Thr	Gln	Lys	Thr	Phe	Ile	Val	Leu
85						90				95					
Asn	Lys	Gly	Lys	Thr	Ile	Phe	Arg	Phe	Ser	Ala	Thr	Asn	Ala	Leu	Tyr
100						105				110					
Val	Leu	Ser	Pro	Phe	His	Pro	Ile	Arg	Arg	Ala	Ala	Val	Lys	Ile	Leu
115						120				125					
Val	His	Ser	Leu	Phe	Asn	Met	Leu	Ile	Met	Cys	Thr	Ile	Leu	Thr	Asn
130						135				140					
Cys	Val	Phe	Met	Ala	Gln	His	Asp	Pro	Pro	Pro	Trp	Thr	Lys	Tyr	Val
145						150				155				160	
Glu	Tyr	Thr	Phe	Thr	Ala	Ile	Tyr	Thr	Phe	Glu	Ser	Leu	Val	Lys	Ile
165						170				175					

-continued

Leu Ala Arg Gly Phe Cys Leu His Ala Phe Thr Phe Leu Arg Asp Pro
 180 185 190
 Trp Asn Trp Leu Asp Phe Ser Val Ile Ile Met Ala Tyr Thr Thr Glu
 195 200 205
 Phe Val Asp Leu Gly Asn Val Ser Ala Leu Arg Thr Phe Arg Val Leu
 210 215 220
 Arg Ala Leu Lys Thr Ile Ser Val Ile Ser Gly Leu Lys Thr Ile Val
 225 230 235 240
 Gly Ala Leu Ile Gln Ser Val Lys Lys Leu Ala Asp Val Met Val Leu
 245 250 255
 Thr Val Phe Cys Leu Ser Val Phe Ala Leu Ile Gly Leu Gln Leu Phe
 260 265 270
 Met Gly Asn Leu Arg His Lys Cys Val Arg Asn Phe Thr Ala Leu Asn
 275 280 285
 Gly Thr Asn Gly Ser Val Glu Ala Asp Gly Leu Val Trp Glu Ser Leu
 290 295 300
 Asp Leu Tyr Leu Ser Asp Pro Glu Asn Tyr Leu Leu Lys Asn Gly Thr
 305 310 315 320
 Ser Asp Val Leu Leu Cys Gly Asn Ser Ser Asp Ala Gly Thr Cys Pro
 325 330 335
 Glu Gly Tyr Arg Cys Leu Lys Ala Gly Glu Asn Pro Asp His Gly Tyr
 340 345 350
 Thr Ser Phe Asp Ser Phe Ala Trp Ala Phe Leu Ala Leu Phe Arg Leu
 355 360 365
 Met Thr Gln Asp Cys Trp Glu Arg Leu Tyr Gln Gln Thr Leu Arg Ser
 370 375 380
 Ala Gly Lys Ile Tyr Met Ile Phe Phe Met Leu Val Ile Phe Leu Gly
 385 390 395 400
 Ser Phe Tyr Leu Val Asn Leu Ile Leu Ala Val Val Ala Met Ala Tyr
 405 410 415
 Glu Glu Gln Asn Gln Ala Thr Ile Ala Glu Thr Glu Glu Lys Glu Lys
 420 425 430
 Arg Phe Gln Glu Ala Met Glu Met Leu Lys Lys Glu His Glu Ala Leu
 435 440 445
 Thr Ile Arg Gly Val Asp Thr Val Ser Arg Ser Ser Leu Glu Met Ser
 450 455 460
 Pro Leu Ala Pro Val Asn Ser His Glu Arg Arg Ser Lys Arg Arg Lys
 465 470 475 480
 Arg Met Ser Ser Gly Thr Glu Glu Cys Gly Glu Asp Arg Leu Pro Lys
 485 490 495
 Ser Asp Ser Glu Asp Gly Pro Arg Ala Met Asn His Leu Ser Leu Thr
 500 505 510
 Arg Gly Leu Ser Arg Thr Ser Met Lys Pro Arg Ser Ser Arg Gly Ser
 515 520 525
 Ile Phe Thr Phe Arg Arg Asp Leu Gly Ser Glu Ala Asp Phe Ala
 530 535 540
 Asp Asp Glu Asn Ser Thr Ala Gly Glu Ser Glu Ser His His Thr Ser
 545 550 555 560
 Leu Leu Val Pro Trp Pro Leu Arg Arg Thr Ser Ala Gln Gly Gln Pro
 565 570

-continued

Ser	Pro	Gly	Thr	Ser	Ala	Pro	Gly	His	Ala	Leu	His	Gly	Lys	Lys	Asn
580						585					590				
Ser	Thr	Val	Asp	Cys	Asn	Gly	Val	Val	Ser	Leu	Leu	Gly	Ala	Gly	Asp
595						600					605				
Pro	Glu	Ala	Thr	Ser	Pro	Gly	Ser	His	Leu	Leu	Arg	Pro	Val	Met	Leu
610						615					620				
Glu	His	Pro	Pro	Asp	Thr	Thr	Thr	Pro	Ser	Glu	Glu	Pro	Gly	Gly	Pro
625						630				635					640
Gln	Met	Leu	Thr	Ser	Gln	Ala	Pro	Cys	Val	Asp	Gly	Phe	Glu	Glu	Pro
645						650				655					
Gly	Ala	Arg	Gln	Arg	Ala	Leu	Ser	Ala	Val	Ser	Val	Leu	Thr	Ser	Ala
660						665				670					
Leu	Glu	Glu	Leu	Glu	Glu	Ser	Arg	His	Lys	Cys	Pro	Pro	Cys	Trp	Asn
675						680				685					
Arg	Leu	Ala	Gln	Arg	Tyr	Leu	Ile	Trp	Glu	Cys	Cys	Pro	Leu	Trp	Met
690						695				700					
Ser	Ile	Lys	Gln	Gly	Val	Lys	Leu	Val	Val	Met	Asp	Pro	Phe	Thr	Asp
705						710				715					720
Leu	Thr	Ile	Thr	Met	Cys	Ile	Val	Leu	Asn	Thr	Leu	Phe	Met	Ala	Leu
725						730				735					
Glu	His	Tyr	Asn	Met	Thr	Ser	Glu	Phe	Glu	Met	Leu	Gln	Val	Gly	
740						745				750					
Asn	Leu	Val	Phe	Thr	Gly	Ile	Phe	Thr	Ala	Glu	Met	Thr	Phe	Lys	Ile
755						760				765					
Ile	Ala	Leu	Asp	Pro	Tyr	Tyr	Phe	Gln	Gln	Gly	Trp	Asn	Ile	Phe	
770						775				780					
Asp	Ser	Ile	Ile	Val	Ile	Leu	Ser	Leu	Met	Glu	Leu	Gly	Leu	Ser	Arg
785						790				795					800
Met	Ser	Asn	Leu	Ser	Val	Leu	Arg	Ser	Phe	Arg	Leu	Leu	Arg	Val	Phe
805						810				815					
Lys	Leu	Ala	Lys	Ser	Trp	Pro	Thr	Leu	Asn	Thr	Leu	Ile	Lys	Ile	Ile
820						825				830					
Gly	Asn	Ser	Val	Gly	Ala	Leu	Gly	Asn	Leu	Thr	Leu	Val	Leu	Ala	Ile
835						840				845					
Ile	Val	Phe	Ile	Phe	Ala	Val	Val	Gly	Met	Gln	Leu	Phe	Gly	Lys	Asn
850						855				860					
Tyr	Ser	Glu	Leu	Arg	Asp	Ser	Asp	Ser	Gly	Leu	Leu	Pro	Arg	Trp	His
865						870				875					880
Met	Met	Asp	Phe	Phe	His	Ala	Phe	Leu	Ile	Ile	Phe	Arg	Ile	Leu	Cys
885						890				895					
Gly	Glu	Trp	Ile	Glu	Thr	Met	Trp	Asp	Cys	Met	Glu	Val	Ser	Gly	Gln
900						905				910					
Ser	Leu	Cys	Leu	Leu	Val	Phe	Leu	Leu	Val	Met	Val	Ile	Gly	Asn	Leu
915						920				925					
Val	Val	Leu	Asn	Leu	Phe	Leu	Ala	Leu	Leu	Leu	Ser	Ser	Phe	Ser	Ala
930						935				940					
Asp	Asn	Leu	Thr	Ala	Pro	Asp	Glu	Asp	Arg	Glu	Met	Asn	Asn	Leu	Gln
945						950				955					960
Leu	Ala	Leu	Ala	Arg	Ile	Gln	Arg	Gly	Leu	Arg	Phe	Val	Lys	Arg	Thr
965						970				975					
Thr	Trp	Asp	Phe	Cys	Cys	Gly	Leu	Leu	Arg	Gln	Arg	Pro	Gln	Lys	Pro

-continued

980	985	990
Ala Ala Leu Ala Ala Gln Gly Gln Leu Pro Ser Cys Ile Ala Thr Pro		
995	1000	1005
Tyr Ser Pro Pro Pro Glu Thr Glu Lys Val Pro Pro Thr Arg Lys		
1010	1015	1020
Glu Thr Arg Phe Glu Glu Gly Glu Gln Pro Gly Gln Gly Thr Pro Gly		
1025	1030	1035
Asp Pro Glu Pro Val Cys Val Pro Ile Ala Val Ala Glu Ser Asp Thr		
1045	1050	1055
Asp Asp Gln Glu Glu Asp Glu Glu Asn Ser Leu Gly Thr Glu Glu Glu		
1060	1065	1070
Ser Ser Lys Gln Glu Ser Gln Pro Val Ser Gly Gly Pro Glu Ala Pro		
1075	1080	1085
Pro Asp Ser Arg Thr Trp Ser Gln Val Ser Ala Thr Ala Ser Ser Glu		
1090	1095	1100
Ala Glu Ala Ser Ala Ser Gln Ala Asp Trp Arg Gln Gln Trp Lys Ala		
1105	1110	1115
1120		
Glu Pro Gln Ala Pro Gly Cys Gly Glu Thr Pro Glu Asp Ser Cys Ser		
1125	1130	1135
Glu Gly Ser Thr Ala Asp Met Thr Asn Thr Ala Glu Leu Leu Glu Gln		
1140	1145	1150
Ile Pro Asp Leu Gly Gln Asp Val Lys Asp Pro Glu Asp Cys Phe Thr		
1155	1160	1165
Glu Gly Cys Val Arg Arg Cys Pro Cys Cys Ala Val Asp Thr Thr Gln		
1170	1175	1180
Ala Pro Gly Lys Val Trp Trp Arg Leu Arg Lys Thr Cys Tyr His Ile		
1185	1190	1195
1200		
Val Glu His Ser Trp Phe Glu Thr Phe Ile Ile Phe Met Ile Leu Leu		
1205	1210	1215
Ser Ser Gly Ala Leu Ala Phe Glu Asp Ile Tyr Leu Glu Glu Arg Lys		
1220	1225	1230
Thr Ile Lys Val Leu Leu Glu Tyr Ala Asp Lys Met Phe Thr Tyr Val		
1235	1240	1245
Phe Val Leu Glu Met Leu Leu Lys Trp Val Ala Tyr Gly Phe Lys Lys		
1250	1255	1260
Tyr Phe Thr Asn Ala Trp Cys Trp Leu Asp Phe Leu Ile Val Asp Val		
1265	1270	1275
1280		
Ser Leu Val Ser Leu Val Ala Asn Thr Leu Gly Phe Ala Glu Met Gly		
1285	1290	1295
Pro Ile Lys Ser Leu Arg Thr Leu Arg Ala Leu Arg Pro Leu Arg Ala		
1300	1305	1310
Leu Ser Arg Phe Glu Gly Met Arg Val Val Val Asn Ala Leu Val Gly		
1315	1320	1325
Ala Ile Pro Ser Ile Met Asn Val Leu Leu Val Cys Leu Ile Phe Trp		
1330	1335	1340
Leu Ile Phe Ser Ile Met Gly Val Asn Leu Phe Ala Gly Lys Phe Gly		
1345	1350	1355
1360		
Arg Cys Ile Asn Gln Thr Glu Gly Asp Leu Pro Leu Asn Tyr Thr Ile		
1365	1370	1375
Val Asn Asn Lys Ser Gln Cys Glu Ser Leu Asn Leu Thr Gly Glu Leu		
1380	1385	1390

-continued

Tyr Trp Thr Lys Val Lys Val Asn Phe Asp Asn Val Gly Ala Gly Tyr
1395 1400 1405

Leu Ala Leu Leu Gln Val Ala Thr Phe Lys Gly Trp Met Asp Ile Met
1410 1415 1420

Tyr Ala Ala Val Asp Ser Arg Gly Tyr Glu Glu Gln Pro Gln Trp Glu
1425 1430 1435 1440

Tyr Asn Leu Tyr Met Tyr Ile Tyr Phe Val Ile Phe Ile Ile Phe Gly
1445 1450 1455

Ser Phe Phe Thr Leu Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe
1460 1465 1470

Asn Gln Gln Lys Lys Lys Leu Gly Gly Gln Asp Ile Phe Met Thr Glu
1475 1480 1485

Glu Gln Lys Lys Tyr Tyr Asn Ala Met Lys Lys Leu Gly Ser Lys Lys
1490 1495 1500

Pro Gln Lys Pro Ile Pro Arg Pro Leu Asn Lys Tyr Gln Gly Phe Ile
1505 1510 1515 1520

Phe Asp Ile Val Thr Lys Gln Ala Phe Asp Val Thr Ile Met Phe Leu
1525 1530 1535

Ile Cys Leu Asn Met Val Thr Met Met Val Glu Thr Asp Asp Gln Ser
1540 1545 1550

Pro Glu Lys Ile Asn Ile Leu Ala Lys Ile Asn Leu Leu Phe Val Ala
1555 1560 1565

Ile Phe Thr Gly Glu Cys Ile Val Lys Leu Ala Ala Leu Arg His Tyr
1570 1575 1580

Tyr Phe Thr Asn Ser Trp Asn Ile Phe Asp Phe Val Val Val Ile Leu
1585 1590 1595 1600

Ser Ile Val Gly Thr Val Leu Ser Asp Ile Ile Gln Lys Tyr Phe Phe
1605 1610 1615

Ser Pro Thr Leu Phe Arg Val Ile Arg Leu Ala Arg Ile Gly Arg Ile
1620 1625 1630

Leu Arg Leu Ile Arg Gly Ala Lys Gly Ile Arg Thr Leu Leu Phe Ala
1635 1640 1645

Leu Met Met Ser Leu Pro Ala Leu Phe Asn Ile Gly Leu Leu Phe
1650 1655 1660

Leu Val Met Phe Ile Tyr Ser Ile Phe Gly Met Ala Asn Phe Ala Tyr
1665 1670 1675 1680

Val Lys Trp Glu Ala Gly Ile Asp Asp Met Phe Asn Phe Gln Thr Phe
1685 1690 1695

Ala Asn Ser Met Leu Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp
1700 1705 1710

Asp Gly Leu Leu Ser Pro Ile Leu Asn Thr Gly Pro Pro Tyr Cys Asp
1715 1720 1725

Pro Thr Leu Pro Asn Ser Asn Gly Ser Arg Gly Asp Cys Gly Ser Pro
1730 1735 1740

Ala Val Gly Ile Leu Phe Phe Thr Thr Tyr Ile Ile Ile Ser Phe Leu
1745 1750 1755 1760

Ile Val Val Asn Met Tyr Ile Ala Ile Ile Leu Glu Asn Phe Ser Val
1765 1770 1775

Ala Thr Glu Glu Ser Thr Glu Pro Leu Ser Glu Asp Asp Phe Asp Met
1780 1785 1790

-continued

Phe	Tyr	Glu	Ile	Trp	Glu	Lys	Phe	Asp	Pro	Glu	Ala	Thr	Gln	Phe	Ile
1795				1800						1805					
Glu	Tyr	Ser	Val	Leu	Ser	Asp	Phe	Ala	Asp	Ala	Leu	Ser	Glu	Pro	Leu
1810				1815							1820				
Arg	Ile	Ala	Lys	Pro	Asn	Gln	Ile	Ser	Leu	Ile	Asn	Met	Asp	Leu	Pro
1825				1830					1835			1840			
Met	Val	Ser	Gly	Asp	Arg	Ile	His	Cys	Met	Asp	Ile	Leu	Phe	Ala	Phe
1845				1850					1855						
Thr	Lys	Arg	Val	Leu	Gly	Glu	Ser	Gly	Glu	Met	Asp	Ala	Leu	Lys	Ile
1860				1865					1870						
Gln	Met	Glu	Glu	Lys	Phe	Met	Ala	Ala	Asn	Pro	Ser	Lys	Ile	Ser	Tyr
1875				1880					1885						
Glu	Pro	Ile	Thr	Thr	Thr	Leu	Arg	Arg	Lys	His	Glu	Glu	Val	Ser	Ala
1890				1895					1900						
Met	Val	Ile	Gln	Arg	Ala	Phe	Arg	Arg	His	Leu	Leu	Gln	Arg	Ser	Leu
1905				1910					1915			1920			
Lys	His	Ala	Ser	Phe	Leu	Phe	Arg	Gln	Gln	Ala	Gly	Ser	Gly	Leu	Ser
1925				1930					1935						
Glu	Glu	Asp	Ala	Pro	Glu	Arg	Glu	Gly	Leu	Ile	Ala	Tyr	Val	Met	Ser
1940				1945					1950						
Glu	Asn	Phe	Ser	Arg	Pro	Leu	Gly	Pro	Pro	Ser	Ser	Ser	Ser	Ile	Ser
1955				1960					1965						
Ser	Thr	Ser	Phe	Pro	Pro	Ser	Tyr	Asp	Ser	Val	Thr	Arg	Ala	Thr	Ser
1970				1975					1980						
Asp	Asn	Leu	Gln	Val	Arg	Gly	Ser	Asp	Tyr	Ser	His	Ser	Glu	Asp	Leu
1985				1990				1995			2000				
Ala	Asp	Phe	Pro	Pro	Ser	Pro	Asp	Arg	Asp	Arg	Glu	Ser	Ile	Val	
2005				2010				2015							

<210> SEQ ID NO 47

<211> LENGTH: 8452

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 47

ggagcccgag	cggagcggcg	gagccgagac	ggcggcggcg	cccgtggat	gtggggatcg	60
cgcggccggg	gcccgtggac	ctgcgcggcg	tgccccaaac	cctacggccg	accgaggccc	120
cgcggcgctg	cagccgcccc	ccccttagggcg	cggggccgggg	actaggcact	tccttccagg	180
cagcctgagg	agagcctgtg	cccccagaag	caggatgaga	agatggcaaa	cttcctgtta	240
cctcgcccc	ccagcagctt	ccgttaggttc	acccggggagt	cactggccgc	catcgagaag	300
cgcacatggctg	aaaaacaacgc	ccgtggttcg	gccaccccac	aggagagccg	tgaggggctg	360
ccagaggagg	agggtcccccg	gccccagctg	gacctacagg	cctccaaaaa	gctgccaat	420
ctctatggca	acccaccccg	agagcttatt	ggggagcccc	tggagacact	ggaccccttc	480
tatagtagcc	agaagacatt	catcgctgtg	aataaggggca	aaaccatctt	ccggttcagt	540
gccaccaatg	ccttgcgtacgt	cctcagcccc	ttccaccccg	tgccgcagac	ggctgtgaag	600
attttggatc	attcgcttt	cagcatgttc	atcatgtgca	ccatcctaacc	caactgtgt	660
ttcatggcac	aacacgaccc	tccgccttgg	accaaataatg	ttgagttacac	cttcacccgc	720
atctacaccc	ttgagttctt	ggtcaagatt	ctagctcgag	gcttctgcct	gcatgogttc	780

-continued

accttcctcc	gggacccgtg	gaactggcta	gacttcagtg	tgatcgtcat	ggcgatgtta	840
tcagagaata	taaagctagg	caatttgcg	gctcttgcga	ctttcagagt	cctgagagct	900
ctgaaaacga	tttcagttat	tccaggcctg	aagacaatcg	tgggagccct	aatccagtct	960
gtgaagaagc	tagccgatgt	gatggtectt	actgtcttct	gcctcagtgt	ctttgcctc	1020
attggcctgc	agctcttcat	gggcaacctg	aggcacaagt	gcgtgegtaa	ctttactgag	1080
ctcaatggca	ccaatggttc	cgtggaggcc	gatggcatag	tctggaaactc	tctggacgtc	1140
tacctcaatg	accaggccaa	ttacctactc	aagaatggca	ccacggatgt	gttgctgtgt	1200
gggaaacagct	ctgatggcgg	gacatgtctt	gaaggctacc	ggtgcctgaa	ggcaggttag	1260
aaccctgacc	acggttacac	cagcttcgac	tccttcgcct	gggctttctt	tgctgttttc	1320
cgcctgatga	cacaggactg	ctgggaacgt	ctataccagc	agacccttag	gtccgcagga	1380
aagatctaca	tgatcttctt	catgcttgc	atattctgg	gtctttctta	ccttgtaaac	1440
ttgatcctgg	ctgtgggtggc	catggcttac	gaggagaaaa	accaaggccac	catcgccgag	1500
acggaagaga	aggagaagcg	tttccaggag	gccatggaga	tgctcaagaa	ggaacacagag	1560
gctctcacca	tcaggggcgt	ggatactgtg	tcccgttagct	ctctggagat	gtctcccttg	1620
gccccagtaa	ccaaccatga	gagaaggagc	aagaggagga	aacgactgtc	ttcaggagaca	1680
gaggatggcg	gagatgacag	gtctcccaag	tcggactcag	aagacggtcc	cagagcattt	1740
aatcagctca	gcctcaccca	tgggctcagc	cgacatcca	tgaggccccg	ctccagccga	1800
gggagcattt	tcacattccg	gcgacgggac	caaggctccg	aggcagattt	tgctagatgc	1860
gagaacagta	cacggggaga	gagcgagagc	cacccgacat	cgctgctgg	gcccctggccc	1920
ctgcgcgcgc	ccagcaccca	agggcagccc	ggctttggaa	cctcagtc	ttgtcatgtc	1980
ctcaatggca	aaagaaacag	caccgtggac	tgcaatgggg	tggttccctt	gctgggggca	2040
ggggacgcag	aggccacctc	tccaggggac	caccttctcc	gccctatagt	gctggacaga	2100
cccccagaca	cgaccacgcc	gtcagaggag	ccgggtgggc	cccagatgt	gacaccgcag	2160
gctccatgtg	cagatggctt	tgaggagccc	ggggcacggc	aacgggcact	cagcgccgtc	2220
agtgtcctca	ccagcgctt	ggaagagttg	gaggagttcc	accgtaaatg	tccaccatgc	2280
tggaaatcgct	ttgcccagca	ttacctcatc	tgggagtgt	gccctacttg	gatgtccatc	2340
aacgagaagg	tgaagttcgt	ggtcatggac	ccgtttgt	acctcaccat	caccatgtgc	2400
atcgtgctca	acacgctt	catggctctg	gagcactaca	acatgacagc	cgagttttag	2460
gagatgctgc	aggtcggaaa	cttggcttc	acgggaatct	tcacggcaga	gatgacccctc	2520
aagatcatcg	cccttgaccc	atactactac	ttccagcagg	gctggatat	cttcgacagc	2580
atcatcgatca	tcctcagect	catggagctg	ggcctgtctc	gcatggccaa	cctgtctgt	2640
ctgcgttct	tccgtctgt	gccccgttcc	aagctggcca	aatcatggcc	caccctgaac	2700
acactcatca	agatcatcg	gaactctgtt	ggtgccctgg	ggaacctgac	cctgggtctg	2760
gccatcatcg	tcttcatctt	cgccgtggtg	ggtatgcagc	tcttcggcaa	gaactactca	2820
gagctgagggc	accgaatcg	cgactctggc	ctgctgcccc	gctggcacat	gatggacttc	2880
ttccacgcct	tcctcatcat	cttccgcate	ctctgtggag	agtggatcga	gaccatgtgg	2940
gactgcatgg	aggtgtctgg	gcagtgcgtg	tgtttgtgtgg	tcttcctgt	ggtcatggtc	3000
attggcaacc	ttgtgggtct	gaatcttcc	ttggccttgc	tgctcagtc	cttcagcgca	3060

-continued

gacaacctca cagccctga cgaggatggg gagatgaaca acctccagct ggccctggct	3120
cgcacccaaa gggccctcg ctttgtcaag cggaccacct gggacttctg ctgcgggctc	3180
ctcgccggc gacctaagaa gcccgcggct cttgccaccc acagccagct gcccaagctgt	3240
atcgccgctc ccagggtcccc accaccccca gaggtggaga aggccccc agcccgcaag	3300
gaaacacggt tcgaggaaga caagcggccg ggccaggggca cccctgggta taccgagcct	3360
gtgtgtgttc ccatacgcaagt ggctgagtca gacactgatg accaggaaga ggatgaggag	3420
aacagcctg gcacggagga ggaagagtcc agcaaacagc aggaatccca agttgtgtct	3480
ggtggcccaeg agccccccca ggagcccaagg ggctggagcc aggtgtcaga gaccacatcc	3540
tctgaagccg aggccagttac gtctcaggca gactggcagc aagageggga agcggagccc	3600
cgggcccccgg ggtgcgggtga gactcccgag gacagttact ccgaggtagt caccgtgac	3660
atgaccaaca ccgcagacact cctggagca atcccaagacc tcggggagga cgtcaaggac	3720
ccagaggact gcttcactga aggctgegtc cgacgttgct cttgctcat ggtggacacg	3780
acgcaggccc cagggaaaggt ctgggtggaga ctgcgcaga a cctgctaccg catagtggag	3840
cacagcttgt tcgagacttt catcatcttc atgatcctgc tcagcagccg agcgtggcc	3900
ttcgaggaca tctaccttga ggagcggaaag accatcaagg ttctgttggaa gtacgcccac	3960
aagatgttca cctatgttct tgtgttggag atgctgtca agtgggtggc ctacggttcc	4020
aagaagtact tcaccaacgc ctgggtctgg ctggacttcc taattgttggaa cgtctcaactg	4080
gtcagccctcg tggccaacac cttaggctt gcccagatgg gtcccatcaa gtccctgagg	4140
acactgcgtg cacttcgacc cctgagggtt ttgtcacgtat tcgaggccat gagggtgggt	4200
gtcaacgcgc tggtggccgc catccccctcc atcatgaatg tcctcctcgt ctgcctcate	4260
ttctggctca tcttcagcat catggcgctg aacctttcg cccggaaatggt cggtaggtgc	4320
atcaaccaga ccgaaggggaa cctgcctctg aactacacca ttgtgaacaa caagagcgc	4380
tgtgagtctt tcaacgtgac tggagagttg tactggacca aggtgaaatg caactttgac	4440
aacgtggggg ctgggtacct ggcctgtctg cagggtggca catttaaagg ctggatggac	4500
atcatgtatg cggctgttga ctccagaggat tatgaagagc agccgcagtg ggaggacaac	4560
ctctacatgt acatctactt cgtcgcttc atcatttcg gtccttctt caccctcaac	4620
ctgtttattt gtgtcatcat tgataacttc aaccagcaga agaaaaatggt agggggccag	4680
gacatcttca tgacggagga gcagaagaag tactacaatg ccatgaagaa gctgggttcc	4740
aagaaacccc agaagcccat cccacggccct ctgaacaatg accagggttt catattcgac	4800
attgtgacca agcaggccct cgtatgttacc atcatgttcc tcatctgttt aaacatggtg	4860
accatgtatgg tagagacaga tgatcagacg cctgaaaaatg tcaacatctt ggccaagatc	4920
aacctgtct tgcgtggccat cttcacagga gagttatgt tcaagatggc tgccctgcgc	4980
cactattact tcaaccaacag ctggaaacatc ttcgatttcg tgggtgtcat cctctccatc	5040
gttggactg tcctctccga catcatccag aagtacttct tctccccaaac gctttccgt	5100
gtcatccgtc tggccaggat cggccgcata ctcaggctga tccgaggagc caaggggatt	5160
cgcacgttgc tcttcgcctt catgtgtcc ctgccccccc tcttcaacat cggccctctc	5220
ctttccctcg tcatgttcat ctactccatc ttcgcatgg ccaactttgc ctacgtcaag	5280
tggaggctg gcatcgacga catgttcaac ttccagaccc ttgccaacag catgtgtgc	5340

-continued

cttttccaga	tcaccacgtc	ggctggctgg	gatggcctcc	tcagccccat	cctcaacaca	5400
ggaccccccct	actgtgaccc	caacctgc	aacagcaatg	gctccagggg	gaactgtggg	5460
agcccagcag	tgggcatect	cttcttcacc	acctacatca	tcatctcctt	cctcatcg	5520
gtcaacatgt	acattgccc	catcctggag	aacttcagt	tggccacaga	agagagcaca	5580
gagccctga	gcgaggatga	cttcgacatg	ttctatgaga	tctggagaa	gttcgacccg	5640
gaggccaccc	agttcattga	gtatggcc	ctgtccgact	ttgccgatgc	cctgtctgag	5700
ccactccgca	tcgccaagcc	caaccagata	agcctcatca	atatggacct	gcccatgg	5760
agcggggacc	gcattccactg	tatggacatc	ctgttcgct	tcaccaagag	ggtgctcggg	5820
gagtctgggg	agatggatgc	cctgaagatc	cagatggagg	agaaattcat	ggcggccaa	5880
ctttccaaga	tctcctatga	geccatcacc	accaccctgc	ggcgaaaaca	tgaggagg	5940
tcggccacgg	tcatccagcg	ggcctccgg	aggcacctcc	tgca	gctc	6000
gcttccttcc	tctccgca	gcaagcgggc	agcagtggcc	tctctgatga	ggatgc	6060
gagcgagagg	gactcattgc	ctacatgatg	aatgagaact	tctccggcg	cagcgg	6120
ctctccagct	cttccatctc	ctccacgtcc	ttcccacccgt	cctatgacag	cgtcacgagg	6180
gccaccagt	ataacctccc	agtgcgtgc	tcggactaca	gccgcagcga	agatcttgc	6240
gactccctc	catctccaga	tagggaccga	gagtctata	tgtgaac	ctg	6300
cggccaggac	gtgcctgtt	gtatcctgac	aaaccta	ag	ttgc	6360
taaggccctc	ctagctctgg	aggtgacagg	ggaac	cttca	gg	6420
gagttctgt	tgtctgtt	gaagctagag	gcggctagag	caacc	cttgc	6480
ctggcacacc	ctagtagccg	ggtctggcca	ggcagcgc	catggctct	aa	6540
ctgtccccc	ctgaggtgaa	acggaaaca	cactgtacat	gttgt	aa	6600
gttttattat	at	ttttactcg	agcaa	aga	ac	6660
gcagcagt	tcacggc	ttccta	acc	tg	ac	6720
cagagcta	gtggatctt	cgcacccaa	gggagggt	tgg	ttc	6780
gaggagccag	gaaggcctaa	gcacttact	cacacacgc	agtgg	ggag	6840
tgtgcacact	tcacacactg	ggcgggaca	taggccc	gg	acc	6900
caa	atgtggt	tggcgg	ctccaaaggc	gagagg	ttt	6960
gggttctt	ctgtgaactc	actctgt	gcc	tgt	cc	7020
tctcagtc	gttcgctt	aggtc	actgc	tgt	aaa	7080
cctggcttcc	cttcc	tgt	aa	gat	ttc	7140
gcagcctc	tcacagcc	gcac	gt	ttc	gg	7200
aggccaagg	aggagaaggc	tgtcccac	ggcc	gtc	ggc	7260
caaagg	ggtctccatt	tcc	ca	gtc	cc	7320
aaggagat	ccactggggc	tcc	ca	ac	gt	7380
tggccctgt	taggagaa	gc	ct	gg	gg	7440
acaaggccat	catccccat	ggg	ca	gg	gg	7500
ctgactatcc	taggcgt	ctg	gaa	gg	aa	7560
ataagtccaa	tcatccccgt	gac	ct	ct	ttt	7620

-continued

atcttttatt	gttagaggacc	agaccatcgc	tgctgaaggg	acaggetggt	ggagectgag	7680
caggcaggggg	cttggctgcc	caagtctagc	tctccagcag	agccaccctc	ggctgggggc	7740
acccaccta	caggccacct	cagtggcgcc	tgcccactgc	gggctggctg	ctcctaccta	7800
cctcgccgag	ctgtcgagg	gctggaeget	cctggcagtg	ctgaaggagg	gggcactgct	7860
ggcaagtcaa	gtattatgtt	tcttcttgc	acctcagttc	ccttggtgcc	agctggctcc	7920
cgcgcaccc	gcccctggcg	gcttatccaa	tgtccccacg	gagtccggtg	ttgggggttt	7980
gtttcttaact	gtcctggta	tgcctctctg	agtggaaatc	ttattttgt	agatctgtac	8040
tttgctctcg	aggcttggag	agatgtgtgt	ccctccctgg	cactcacttc	aggccgcatg	8100
ggcaggaatg	tggcctgggg	ctggcagggt	tggcagggt	gccatcccag	ctggcggaa	8160
agaggctgtg	gtttcgaata	tgcagatggc	cttggagcct	taagacaggt	gcggggctgg	8220
ctgcagacag	ggtggctgag	gggagggggca	atcgagtccc	gagtcttgg	cttagatgg	8280
cctccttttc	tagaactgca	cagagggggag	gagagagtag	ccggtagacg	gaaggaggga	8340
ggggcatgct	tacatgtcag	gtgttgtttt	aggtaaatga	gcacttctca	ctaagcttcg	8400
tgtataaatg	gaatacatta	tttttcaaac	aaacgaataa	atggcttaca	cg	8452

<210> SEQ ID NO 48

<211> LENGTH: 2020

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 48

Met	Ala	Asn	Phe	Leu	Leu	Pro	Arg	Gly	Thr	Ser	Ser	Phe	Arg	Arg	Phe
1				5					10				15		
Thr	Arg	Glu	Ser	Leu	Ala	Ala	Ile	Glu	Lys	Arg	Met	Ala	Glu	Lys	Gln
20					25				30						
Ala	Arg	Gly	Ser	Ala	Thr	Ser	Gln	Glu	Ser	Arg	Glu	Gly	Leu	Pro	Glu
35					40				45						
Glu	Glu	Ala	Pro	Arg	Pro	Gln	Leu	Asp	Leu	Gln	Ala	Ser	Lys	Lys	Leu
50					55				60						
Pro	Asp	Leu	Tyr	Gly	Asn	Pro	Pro	Arg	Glu	Leu	Ile	Gly	Glu	Pro	Leu
65					70				75				80		
Glu	Asp	Leu	Asp	Pro	Phe	Tyr	Ser	Thr	Gln	Lys	Thr	Phe	Ile	Val	Leu
85					90				95						
Asn	Lys	Gly	Lys	Thr	Ile	Phe	Arg	Phe	Ser	Ala	Thr	Asn	Ala	Leu	Tyr
100					105				110						
Val	Leu	Ser	Pro	Phe	His	Pro	Val	Arg	Arg	Ala	Ala	Val	Lys	Ile	Leu
115					120				125						
Val	His	Ser	Leu	Phe	Ser	Met	Leu	Ile	Met	Cys	Thr	Ile	Leu	Thr	Asn
130					135				140						
Cys	Val	Phe	Met	Ala	Gln	His	Asp	Pro	Pro	Pro	Trp	Thr	Lys	Tyr	Val
145					150				155				160		
Glu	Tyr	Thr	Phe	Thr	Ala	Ile	Tyr	Thr	Phe	Glu	Ser	Leu	Val	Lys	Ile
165					170				175						
Leu	Ala	Arg	Gly	Phe	Cys	Leu	His	Ala	Phe	Thr	Phe	Leu	Arg	Asp	Pro
180					185				190						
Trp	Asn	Trp	Leu	Asp	Phe	Ser	Val	Ile	Val	Met	Ala	Tyr	Val	Ser	Glu
195					200				205						
Asn	Ile	Lys	Leu	Gly	Asn	Leu	Ser	Ala	Leu	Arg	Thr	Phe	Arg	Val	Leu

-continued

210	215	220
Arg Ala Leu Lys Thr Ile Ser Val Ile Pro Gly Leu Lys Thr Ile Val		
225	230	235
Gly Ala Leu Ile Gln Ser Val Lys Lys Leu Ala Asp Val Met Val Leu		
245	250	255
Thr Val Phe Cys Leu Ser Val Phe Ala Leu Ile Gly Leu Gln Leu Phe		
260	265	270
Met Gly Asn Leu Arg His Lys Cys Val Arg Asn Phe Thr Glu Leu Asn		
275	280	285
Gly Thr Asn Gly Ser Val Glu Ala Asp Gly Ile Val Trp Asn Ser Leu		
290	295	300
Asp Val Tyr Leu Asn Asp Pro Ala Asn Tyr Leu Leu Lys Asn Gly Thr		
305	310	315
Thr Asp Val Leu Leu Cys Gly Asn Ser Ser Asp Ala Gly Thr Cys Pro		
325	330	335
Glu Gly Tyr Arg Cys Leu Lys Ala Gly Glu Asn Pro Asp His Gly Tyr		
340	345	350
Thr Ser Phe Asp Ser Phe Ala Trp Ala Phe Leu Ala Leu Phe Arg Leu		
355	360	365
Met Thr Gln Asp Cys Trp Glu Arg Leu Tyr Gln Gln Thr Leu Arg Ser		
370	375	380
Ala Gly Lys Ile Tyr Met Ile Phe Phe Met Leu Val Ile Phe Leu Gly		
385	390	395
Ser Phe Tyr Leu Val Asn Leu Ile Leu Ala Val Val Ala Met Ala Tyr		
405	410	415
Glu Glu Gln Asn Gln Ala Thr Ile Ala Glu Thr Glu Glu Lys Glu Lys		
420	425	430
Arg Phe Gln Glu Ala Met Glu Met Leu Lys Lys Glu His Glu Ala Leu		
435	440	445
Thr Ile Arg Gly Val Asp Thr Val Ser Arg Ser Ser Leu Glu Met Ser		
450	455	460
Pro Leu Ala Pro Val Thr Asn His Glu Arg Arg Ser Lys Arg Arg Lys		
465	470	475
Arg Leu Ser Ser Gly Thr Glu Asp Gly Gly Asp Asp Arg Leu Pro Lys		
485	490	495
Ser Asp Ser Glu Asp Gly Pro Arg Ala Leu Asn Gln Leu Ser Leu Thr		
500	505	510
His Gly Leu Ser Arg Thr Ser Met Arg Pro Arg Ser Ser Arg Gly Ser		
515	520	525
Ile Phe Thr Phe Arg Arg Arg Asp Gln Gly Ser Glu Ala Asp Phe Ala		
530	535	540
Asp Asp Glu Asn Ser Thr Ala Gly Glu Ser Glu Ser His Arg Thr Ser		
545	550	555
Leu Leu Val Pro Trp Pro Leu Arg Arg Pro Ser Thr Gln Gly Gln Pro		
565	570	575
Gly Phe Gly Thr Ser Ala Pro Gly His Val Leu Asn Gly Lys Arg Asn		
580	585	590
Ser Thr Val Asp Cys Asn Gly Val Val Ser Leu Leu Gly Ala Gly Asp		
595	600	605
Ala Glu Ala Thr Ser Pro Gly Ser His Leu Leu Arg Pro Ile Val Leu		
610	615	620

-continued

Asp	Arg	Pro	Pro	Asp	Thr	Thr	Thr	Pro	Ser	Glu	Glu	Pro	Gly	Gly	Pro
625					630				635						640
Gln	Met	Leu	Thr	Pro	Gln	Ala	Pro	Cys	Ala	Asp	Gly	Phe	Glu	Glu	Pro
645					650				655						
Gly	Ala	Arg	Gln	Arg	Ala	Leu	Ser	Ala	Val	Ser	Val	Leu	Thr	Ser	Ala
660					665				670						
Leu	Glu	Glu	Leu	Glu	Glu	Ser	His	Arg	Lys	Cys	Pro	Pro	Cys	Trp	Asn
675					680				685						
Arg	Phe	Ala	Gln	His	Tyr	Leu	Ile	Trp	Glu	Cys	Cys	Pro	Leu	Trp	Met
690					695				700						
Ser	Ile	Lys	Gln	Lys	Val	Lys	Phe	Val	Val	Met	Asp	Pro	Phe	Ala	Asp
705					710				715						720
Leu	Thr	Ile	Thr	Met	Cys	Ile	Val	Leu	Asn	Thr	Leu	Phe	Met	Ala	Leu
725					730				735						
Glu	His	Tyr	Asn	Met	Thr	Ala	Glu	Phe	Glu	Glu	Met	Leu	Gln	Val	Gly
740					745				750						
Asn	Leu	Val	Phe	Thr	Gly	Ile	Phe	Thr	Ala	Glu	Met	Thr	Phe	Lys	Ile
755					760				765						
Ile	Ala	Leu	Asp	Pro	Tyr	Tyr	Tyr	Phe	Gln	Gln	Gly	Trp	Asn	Ile	Phe
770					775				780						
Asp	Ser	Ile	Ile	Val	Ile	Leu	Ser	Leu	Met	Glu	Leu	Gly	Leu	Ser	Arg
785					790				795						800
Met	Gly	Asn	Leu	Ser	Val	Leu	Arg	Ser	Phe	Arg	Leu	Leu	Arg	Val	Phe
805					810				815						
Lys	Leu	Ala	Lys	Ser	Trp	Pro	Thr	Leu	Asn	Thr	Leu	Ile	Lys	Ile	Ile
820					825				830						
Gly	Asn	Ser	Val	Gly	Ala	Leu	Gly	Asn	Leu	Thr	Leu	Val	Leu	Ala	Ile
835					840				845						
Ile	Val	Phe	Ile	Phe	Ala	Val	Val	Gly	Met	Gln	Leu	Phe	Gly	Lys	Asn
850					855				860						
Tyr	Ser	Glu	Leu	Arg	His	Arg	Ile	Ser	Asp	Ser	Gly	Leu	Leu	Pro	Arg
865					870				875						880
Trp	His	Met	Met	Asp	Phe	Phe	His	Ala	Phe	Leu	Ile	Ile	Phe	Arg	Ile
885					890				895						
Leu	Cys	Gly	Glu	Trp	Ile	Glu	Thr	Met	Trp	Asp	Cys	Met	Glu	Val	Ser
900					905				910						
Gly	Gln	Ser	Leu	Cys	Leu	Leu	Val	Phe	Leu	Leu	Val	Met	Val	Ile	Gly
915					920				925						
Asn	Leu	Val	Val	Leu	Asn	Leu	Phe	Leu	Ala	Leu	Leu	Ser	Ser	Phe	
930					935				940						
Ser	Ala	Asp	Asn	Leu	Thr	Ala	Pro	Asp	Glu	Asp	Gly	Glu	Met	Asn	Asn
945					950				955						960
Leu	Gln	Leu	Ala	Leu	Ala	Arg	Ile	Gln	Arg	Gly	Leu	Arg	Phe	Val	Lys
965					970				975						
Arg	Thr	Thr	Trp	Asp	Phe	Cys	Cys	Gly	Leu	Leu	Arg	Arg	Arg	Pro	Lys
980					985				990						
Lys	Pro	Ala	Ala	Leu	Ala	Thr	His	Ser	Gln	Leu	Pro	Ser	Cys	Ile	Ala
995					1000				1005						
Ala	Pro	Arg	Ser	Pro	Pro	Pro	Pro	Glu	Val	Glu	Lys	Ala	Pro	Pro	Ala
1010					1015				1020						

-continued

Arg Lys Glu Thr Arg Phe Glu Glu Asp Lys Arg Pro Gly Gln Gly Thr			
1025	1030	1035	1040
Pro Gly Asp Thr Glu Pro Val Cys Val Pro Ile Ala Val Ala Glu Ser			
1045	1050	1055	
Asp Thr Asp Asp Gln Glu Glu Asp Glu Glu Asn Ser Leu Gly Thr Glu			
1060	1065	1070	
Glu Glu Glu Ser Ser Lys Gln Gln Glu Ser Gln Val Val Ser Gly Gly			
1075	1080	1085	
His Glu Pro Pro Gln Glu Pro Arg Ala Trp Ser Gln Val Ser Glu Thr			
1090	1095	1100	
Thr Ser Ser Glu Ala Glu Ala Ser Thr Ser Gln Ala Asp Trp Gln Gln			
1105	1110	1115	1120
Glu Arg Glu Ala Glu Pro Arg Ala Pro Gly Cys Gly Glu Thr Pro Glu			
1125	1130	1135	
Asp Ser Tyr Ser Glu Gly Ser Thr Ala Asp Met Thr Asn Thr Ala Asp			
1140	1145	1150	
Leu Leu Glu Gln Ile Pro Asp Leu Gly Glu Asp Val Lys Asp Pro Glu			
1155	1160	1165	
Asp Cys Phe Thr Glu Gly Cys Val Arg Arg Cys Pro Cys Cys Met Val			
1170	1175	1180	
Asp Thr Thr Gln Ala Pro Gly Lys Val Trp Trp Arg Leu Arg Lys Thr			
1185	1190	1195	1200
Cys Tyr Arg Ile Val Glu His Ser Trp Phe Glu Thr Phe Ile Ile Phe			
1205	1210	1215	
Met Ile Leu Leu Ser Ser Gly Ala Leu Ala Phe Glu Asp Ile Tyr Leu			
1220	1225	1230	
Glu Glu Arg Lys Thr Ile Lys Val Leu Leu Glu Tyr Ala Asp Lys Met			
1235	1240	1245	
Phe Thr Tyr Val Phe Val Leu Glu Met Leu Leu Lys Trp Val Ala Tyr			
1250	1255	1260	
Gly Phe Lys Lys Tyr Phe Thr Asn Ala Trp Cys Trp Leu Asp Phe Leu			
1265	1270	1275	1280
Ile Val Asp Val Ser Leu Val Ser Leu Val Ala Asn Thr Leu Gly Phe			
1285	1290	1295	
Ala Glu Met Gly Pro Ile Lys Ser Leu Arg Thr Leu Arg Ala Leu Arg			
1300	1305	1310	
Pro Leu Arg Ala Leu Ser Arg Phe Glu Gly Met Arg Val Val Val Asn			
1315	1320	1325	
Ala Leu Val Gly Ala Ile Pro Ser Ile Met Asn Val Leu Leu Val Cys			
1330	1335	1340	
Leu Ile Phe Trp Leu Ile Phe Ser Ile Met Gly Val Asn Leu Phe Ala			
1345	1350	1355	1360
Gly Lys Phe Gly Arg Cys Ile Asn Gln Thr Glu Gly Asp Leu Pro Leu			
1365	1370	1375	
Asn Tyr Thr Ile Val Asn Asn Lys Ser Glu Cys Glu Ser Phe Asn Val			
1380	1385	1390	
Thr Gly Glu Leu Tyr Trp Thr Lys Val Lys Val Asn Phe Asp Asn Val			
1395	1400	1405	
Gly Ala Gly Tyr Leu Ala Leu Leu Gln Val Ala Thr Phe Lys Gly Trp			
1410	1415	1420	
Met Asp Ile Met Tyr Ala Ala Val Asp Ser Arg Gly Tyr Glu Glu Gln			

-continued

1425	1430	1435	1440
Pro Gln Trp Glu Asp Asn Leu Tyr Met Tyr Ile Tyr Phe Val Val Phe			
1445	1450	1455	
Ile Ile Phe Gly Ser Phe Phe Thr Leu Asn Leu Phe Ile Gly Val Ile			
1460	1465	1470	
Ile Asp Asn Phe Asn Gln Gln Lys Lys Lys Leu Gly Gly Gln Asp Ile			
1475	1480	1485	
Phe Met Thr Glu Glu Gln Lys Lys Tyr Tyr Asn Ala Met Lys Lys Leu			
1490	1495	1500	
Gly Ser Lys Lys Pro Gln Lys Pro Ile Pro Arg Pro Leu Asn Lys Tyr			
1505	1510	1515	1520
Gln Gly Phe Ile Phe Asp Ile Val Thr Lys Gln Ala Phe Asp Val Thr			
1525	1530	1535	
Ile Met Phe Leu Ile Cys Leu Asn Met Val Thr Met Met Val Glu Thr			
1540	1545	1550	
Asp Asp Gln Ser Pro Glu Lys Val Asn Ile Leu Ala Lys Ile Asn Leu			
1555	1560	1565	
Leu Phe Val Ala Ile Phe Thr Gly Glu Cys Ile Val Lys Met Ala Ala			
1570	1575	1580	
Leu Arg His Tyr Tyr Phe Thr Asn Ser Trp Asn Ile Phe Asp Phe Val			
1585	1590	1595	1600
Val Val Ile Leu Ser Ile Val Gly Thr Val Leu Ser Asp Ile Ile Gln			
1605	1610	1615	
Lys Tyr Phe Phe Ser Pro Thr Leu Phe Arg Val Ile Arg Leu Ala Arg			
1620	1625	1630	
Ile Gly Arg Ile Leu Arg Leu Ile Arg Gly Ala Lys Gly Ile Arg Thr			
1635	1640	1645	
Leu Leu Phe Ala Leu Met Met Ser Leu Pro Ala Leu Phe Asn Ile Gly			
1650	1655	1660	
Leu Leu Leu Phe Leu Val Met Phe Ile Tyr Ser Ile Phe Gly Met Ala			
1665	1670	1675	1680
Asn Phe Ala Tyr Val Lys Trp Glu Ala Gly Ile Asp Asp Met Phe Asn			
1685	1690	1695	
Phe Gln Thr Phe Ala Asn Ser Met Leu Cys Leu Phe Gln Ile Thr Thr			
1700	1705	1710	
Ser Ala Gly Trp Asp Gly Leu Leu Ser Pro Ile Leu Asn Thr Gly Pro			
1715	1720	1725	
Pro Tyr Cys Asp Pro Asn Leu Pro Asn Ser Asn Gly Ser Arg Gly Asn			
1730	1735	1740	
Cys Gly Ser Pro Ala Val Gly Ile Leu Phe Phe Thr Thr Tyr Ile Ile			
1745	1750	1755	1760
Ile Ser Phe Leu Ile Val Val Asn Met Tyr Ile Ala Ile Ile Leu Glu			
1765	1770	1775	
Asn Phe Ser Val Ala Thr Glu Glu Ser Thr Glu Pro Leu Ser Glu Asp			
1780	1785	1790	
Asp Phe Asp Met Phe Tyr Glu Ile Trp Glu Lys Phe Asp Pro Glu Ala			
1795	1800	1805	
Thr Gln Phe Ile Glu Tyr Leu Ala Leu Ser Asp Phe Ala Asp Ala Leu			
1810	1815	1820	
Ser Glu Pro Leu Arg Ile Ala Lys Pro Asn Gln Ile Ser Leu Ile Asn			
1825	1830	1835	1840

-continued

Met Asp Leu Pro Met Val Ser Gly Asp Arg Ile His Cys Met Asp Ile
1845 1850 1855

Leu Phe Ala Phe Thr Lys Arg Val Leu Gly Glu Ser Gly Glu Met Asp
1860 1865 1870

Ala Leu Lys Ile Gln Met Glu Glu Lys Phe Met Ala Ala Asn Pro Ser
1875 1880 1885

Lys Ile Ser Tyr Glu Pro Ile Thr Thr Leu Arg Arg Lys His Glu
1890 1895 1900

Glu Val Ser Ala Thr Val Ile Gln Arg Ala Phe Arg Arg His Leu Leu
1905 1910 1915 1920

Gln Arg Ser Val Lys His Ala Ser Phe Leu Phe Arg Gln Gln Ala Gly
1925 1930 1935

Ser Ser Gly Leu Ser Asp Glu Asp Ala Pro Glu Arg Glu Gly Leu Ile
1940 1945 1950

Ala Tyr Met Met Asn Glu Asn Phe Ser Arg Arg Ser Gly Pro Leu Ser
1955 1960 1965

Ser Ser Ser Ile Ser Ser Thr Ser Phe Pro Pro Ser Tyr Asp Ser Val
1970 1975 1980

Thr Arg Ala Thr Ser Asp Asn Leu Pro Val Arg Ala Ser Asp Tyr Ser
1985 1990 1995 2000

Arg Ser Glu Asp Leu Ala Asp Phe Pro Pro Ser Pro Asp Arg Asp Arg
2005 2010 2015

Glu Ser Ile Val
2020

<210> SEQ ID NO 49

<211> LENGTH: 4516

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

ctcggcgccgc	ctggacccct	gcccctctct	gggtggagaa	gctccggcc	gcttccgggt	60
ttaactcctt	tcagcctgg	gctcccaggc	cctctctctt	tttccctggac	tggctctcac	120
ccccttcggt	ccccttcctt	tagctcaggc	tcccttaccc	ttccttttagc	ccacagccca	180
gagtccca	tcctcagtca	ctttcctca	ccaaagggtcc	cagccttct	tcttccccc	240
ctttgcacta	tccctatct	gccccttct	ctatccctag	ggctcagtt	cccacatccg	300
tcctccccc	tcccaggccc	ggagttccag	accttttgtt	ctcctttcg	ggtcgttct	360
gggtccttgc	cccctttccc	cacttttgag	ttccagatgt	caaaccaggc	ctccctccac	420
ccccagaaaa	ttgcttccat	ggaaatgcct	ctctaaaaca	tgaacttttc	ctagagacta	480
cggcagtctc	tcttcccact	tgctgaccct	ttgctaccta	tgtgccgggt	tttactctca	540
tttgggtaag	gtcgaggctg	gtcttggaa	cagcaccatg	gttctgggt	ctggcatctg	600
tggcctctt	ccacatcgga	tcttcccttc	cttactcg	gtgggtgtt	tgggtgggt	660
gctgcctgtt	ctcaggagcc	atggcctcca	gctcagecca	actgccagca	ccattcgaag	720
ctcagagcca	ccacgagaac	gctcgattgg	ggatgtcacc	accgctccac	cgaggtcac	780
cccagagac	cggccctgtta	atcattccgt	cactgtatcat	ggcatgaagc	cgcgcaaggc	840
cttccagtc	ctgggcatcg	actacacaca	cgtgcgcacc	cccttcgaga	tctccctctg	900
gatccttctg	gcctgcctca	tgaagatagg	tttccatgtg	atccccacta	tctcaagcat	960

-continued

cgtccccggag	agctgcctgc	tgatcggtt	ggggctgtcg	gtggggggcc	tgatcaaggg	1020
tgttaggcgag	acacccccc	tcctgcagtc	cgacgtctc	ttccttcc	tgctgccgc	1080
catcatcctg	gatgcgggct	acttcctgcc	actgcggcag	ttcacagaaa	acctgggcac	1140
catcctgate	tttgcgtgg	tgggcacgct	gtggAACGCC	ttcttcctgg	gcggccat	1200
gtacgccgtg	tgccctgggtgg	gccccgtggaa	gatcaacaac	atcgccctcc	tggacaacct	1260
gctttcggc	agcatcatct	cggccgtggaa	ccccgtggcg	gttctggctg	tctttgagga	1320
aattcacatc	aatgagctgc	tgcacatct	tgttttggg	gagtccctgc	tcaatgacgc	1380
cgtcaactgtg	gtccctgtatc	acccctttga	ggagtttgc	aactacgaaac	acgtgggcatt	1440
cgtggacatc	ttccctcggt	tcctgagctt	tttcgtggtg	gccttggcg	gggtgtttgt	1500
gggcgtggtc	tacggggta	tcgcagcctt	caccccccga	tttacctccc	acatccgggt	1560
catcgagccg	ctccctcgat	tcctctacag	ctacatggcc	tacttgtcag	ccgagcttt	1620
ccacctgtca	ggcatcatgg	cgctcatagc	ctcaggagtg	gtgatgcgc	cctatgtgga	1680
ggccaacatc	tcccacaagt	cccacaccac	catcaaatac	ttcctgaaga	tgtggagcag	1740
cgtcagcgag	accctcatct	tcatcttctc	cggcgtctcc	acgggtggcg	gtccccacca	1800
ctggaaactgg	acccctcgta	tcagcaccc	gtctttctgc	ctcatcgccc	gcgtgtggg	1860
ggtgctggc	ctgacctgg	tcatcaacaa	gttccgtatc	gtgaagctga	cccccaagga	1920
ccagttcata	atcgccctatg	ggggcctgcg	agggggccatc	gccttctctc	tgggttaccc	1980
cctggacaag	aaggacttcc	ccatgtgtga	cctgttctc	actgcccata	tcactgtcat	2040
cttcttcacc	gtctttgtgc	agggcatgac	cattcgccc	ctggtagacc	tgttggctgt	2100
gaagaaaaaaag	caagagacga	agcgctcat	caacgaagag	atccacacac	agttcttgaa	2160
ccacccctcg	acaggcatcg	aagacatctg	tggccactac	ggtcaccacc	actggaaagga	2220
caagctcaac	cggtttaata	agaaatatgt	gaagaagtgt	ctgatagctg	gcgagcgctc	2280
caaggagccc	cagctcattt	ccttctacca	caagatggag	atgaagcagg	ccatcgagct	2340
ggtggagagc	gggggcattgg	gcaagatccc	ctctgcgtc	tccacccgt	ccatgcagaa	2400
catccacccc	aagtccctgc	cttccgagcg	catcctgc	gcactgtcca	aggacaagga	2460
ggaggagatc	cgcggaaatcc	tgaggaacaa	cttgcagaag	accaggcagc	ggctgcgggtc	2520
ctacaacaga	cacacgctgg	tggcagaccc	ctacgaggaa	gcctggaaacc	agatgtgt	2580
ccggaggcag	aaggccccggc	agctggagca	gaagatcaac	aactacctga	cggtgcagc	2640
ccacaagctg	gactcacca	ccatgtctcg	ggcccgatc	ggctcagacc	caactggccta	2700
tgagccgaag	gaggacctgc	ctgtcatc	catcgacc	gttccccgc	agtcacccga	2760
gtctgtggac	ctggtaatg	aggagctgaa	ggccaaatgc	ttagggttga	gcggggatcc	2820
tgcaaaagg	gctgaggagg	acgaggacg	cgatggggc	atcatgtgc	ggagcaagga	2880
gacttcgtcc	ccaggaaccc	acgatgtctt	caccccg	cccagtgaca	gcggccatc	2940
ccagaggata	cagcgctcc	tcaatgtacc	aggccccacac	cctgagctg	gggggggaga	3000
accgttcttc	cccaaggggc	agtaacgc	ggccagcag	gcagcgctg	tccctcaca	3060
gactcttcca	ccagagcagg	ggctgctggg	ggctccctt	gccttcctg	acccggattt	3120
gcctgcccc	tccccctacc	gcatggc	tggccacaca	gccccccaccc	cagcacagct	3180
cctccctgc	cgcctcccg	gaagcatct	ccccaccaga	gtgcctccc	caatccattt	3240

-continued

ggcagaactg ctggggctgg tgaggccggc cctgccccctc cctagatcca ggcttctccc	3300
ggacctggac tagggcctcg gaggctcctc cctctgcctc atcctcctcc tcattcagac	3360
caatcttagt ttctaaccas agagtctctg gctcagctgt ggtcccaccc aggaaggag	3420
ggagctgagg cttcccttga gtaggccctg ctttatcagg ggacaaacca ggggtaccag	3480
gcacatggct gggggaggga ctgctgaccc accaaggctc cacactcctc ctgccagctc	3540
tgtcaccctg gccaccaccc aacctatcct tactcagagc tgcgggctga gggcatctct	3600
gagtgtctct gcctggagca ggggtggttt ctacggtgac agtgcacgtga ctcagagctt	3660
ttcgaactgt gctcccacgg ggaccactgg gcccctcagg ggaagctgct aggggaagga	3720
ctggccgtgg ctccagaatg tgctgcctt ttaagtttg tttgttcaca ctccatatata	3780
tgattgtttt cacagagggc gtcctgtttt taaaacactt ttgaaaaccc ctggctgaac	3840
agtgtctgc ctctaactcc ctccctcacac tccagaatta cccttcctca tctgtgcctg	3900
tctgtccaaac ccctccccca cgtctctctg cctgctggc tcttaactgt tgctcagaaga	3960
ctgtgacatc agaagtaact cccactccta atcaagagtc tctccagcct cacagatgt	4020
ggcctcttgg cacctgccta gtcctggcctc gtcctgtgg cctgctctta	4080
cttccccac cctgggtttt gcccctggaa ctttcctt gtgtgtacca caccctgcct	4140
gctgtggagc ccattgtgga ggccgtgggg gggagaaggc ctcctctgag gateccctgt	4200
ccctggggc tggtgattt ggcagaatcc tgggccccca gagacctttt cccacacaca	4260
ctcctcccccc tttccctgg ggcactcccccc caggatttgtt caatagtcag agtgcctt	4320
tttgcagggg actgggccat gggcctcgg cccatctgtc catcctcctc tccatgcaag	4380
tgctgtttgg gcaggagtca ccattgttgcgtt gtcacatcga caaccacgtt ccaagccacc	4440
gcagctgtg ccactctgtt gtcgttacag aagaaactga atcttttca tattctaata	4500
aatcaatgtg agtttt	4516

<210> SEQ ID NO 50
<211> LENGTH: 815
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 50

Met Val Leu Arg Ser Gly Ile Cys Gly Leu Ser Pro His Arg Ile Phe			
1	5	10	15
Pro Ser Leu Leu Val Val Val Ala Leu Val Gly Leu Leu Pro Val Leu			
20	25	30	
Arg Ser His Gly Leu Gln Leu Ser Pro Thr Ala Ser Thr Ile Arg Ser			
35	40	45	
Ser Glu Pro Pro Arg Glu Arg Ser Ile Gly Asp Val Thr Thr Ala Pro			
50	55	60	
Pro Glu Val Thr Pro Glu Ser Arg Pro Val Asn His Ser Val Thr Asp			
65	70	75	80
His Gly Met Lys Pro Arg Lys Ala Phe Pro Val Leu Gly Ile Asp Tyr			
85	90	95	
Thr His Val Arg Thr Pro Phe Glu Ile Ser Leu Trp Ile Leu Leu Ala			
100	105	110	
Cys Leu Met Lys Ile Gly Phe His Val Ile Pro Thr Ile Ser Ser Ile			
115	120	125	

-continued

Val	Pro	Glu	Ser	Cys	Leu	Leu	Ile	Val	Val	Gly	Leu	Leu	Val	Gly	Gly
130				135						140					
Leu	Ile	Lys	Gly	Val	Gly	Glu	Thr	Pro	Pro	Phe	Leu	Gln	Ser	Asp	Val
145				150						155					160
Phe	Phe	Leu	Phe	Leu	Leu	Pro	Pro	Ile	Ile	Leu	Asp	Ala	Gly	Tyr	Phe
165				170						175					
Leu	Pro	Leu	Arg	Gln	Phe	Thr	Glu	Asn	Leu	Gly	Thr	Ile	Leu	Ile	Phe
180				185					190						
Ala	Val	Val	Gly	Thr	Leu	Trp	Asn	Ala	Phe	Phe	Leu	Gly	Gly	Leu	Met
195				200				205							
Tyr	Ala	Val	Cys	Leu	Val	Gly	Gly	Glu	Gln	Ile	Asn	Asn	Ile	Gly	Leu
210				215					220						
Leu	Asp	Asn	Leu	Leu	Phe	Gly	Ser	Ile	Ile	Ser	Ala	Val	Asp	Pro	Val
225				230				235		240					
Ala	Val	Leu	Ala	Val	Phe	Glu	Glu	Ile	His	Ile	Asn	Glu	Leu	Leu	His
245				250				255							
Ile	Leu	Val	Phe	Gly	Glu	Ser	Leu	Leu	Asn	Asp	Ala	Val	Thr	Val	Val
260				265				270							
Leu	Tyr	His	Leu	Phe	Glu	Glu	Phe	Ala	Asn	Tyr	Glu	His	Val	Gly	Ile
275				280				285							
Val	Asp	Ile	Phe	Leu	Gly	Phe	Leu	Ser	Phe	Phe	Val	Val	Ala	Leu	Gly
290				295				300							
Gly	Val	Leu	Val	Gly	Val	Val	Tyr	Gly	Val	Ile	Ala	Ala	Phe	Thr	Ser
305				310				315		320					
Arg	Phe	Thr	Ser	His	Ile	Arg	Val	Ile	Glu	Pro	Leu	Phe	Val	Phe	Leu
325				330				335							
Tyr	Ser	Tyr	Met	Ala	Tyr	Leu	Ser	Ala	Glu	Leu	Phe	His	Leu	Ser	Gly
340				345				350							
Ile	Met	Ala	Leu	Ile	Ala	Ser	Gly	Val	Val	Met	Arg	Pro	Tyr	Val	Glu
355				360				365							
Ala	Asn	Ile	Ser	His	Lys	Ser	His	Thr	Thr	Ile	Lys	Tyr	Phe	Leu	Lys
370				375				380							
Met	Trp	Ser	Ser	Val	Ser	Glu	Thr	Leu	Ile	Phe	Ile	Phe	Leu	Gly	Val
385				390				395		400					
Ser	Thr	Val	Ala	Gly	Ser	His	His	Trp	Asn	Trp	Thr	Phe	Val	Ile	Ser
405				410				415							
Thr	Leu	Leu	Phe	Cys	Leu	Ile	Ala	Arg	Val	Leu	Gly	Val	Leu	Gly	Leu
420				425				430							
Thr	Trp	Phe	Ile	Asn	Lys	Phe	Arg	Ile	Val	Lys	Leu	Thr	Pro	Lys	Asp
435				440				445							
Gln	Phe	Ile	Ile	Ala	Tyr	Gly	Gly	Leu	Arg	Gly	Ala	Ile	Ala	Phe	Ser
450				455				460							
Leu	Gly	Tyr	Leu	Leu	Asp	Lys	Lys	His	Phe	Pro	Met	Cys	Asp	Leu	Phe
465				470				475		480					
Leu	Thr	Ala	Ile	Ile	Thr	Val	Ile	Phe	Phe	Thr	Val	Phe	Val	Gln	Gly
485				490				495							
Met	Thr	Ile	Arg	Pro	Leu	Val	Asp	Leu	Leu	Ala	Val	Lys	Lys	Lys	Gln
500				505				510							
Glu	Thr	Lys	Arg	Ser	Ile	Asn	Glu	Glu	Ile	His	Thr	Gln	Phe	Leu	Asp
515				520				525							

-continued

His	Leu	Leu	Thr	Gly	Ile	Glu	Asp	Ile	Cys	Gly	His	Tyr	Gly	His	His
530					535				540						
His	Trp	Lys	Asp	Lys	Leu	Asn	Arg	Phe	Asn	Lys	Lys	Tyr	Val	Lys	Lys
545					550				555				560		
Cys	Leu	Ile	Ala	Gly	Glu	Arg	Ser	Lys	Glu	Pro	Gln	Leu	Ile	Ala	Phe
565					570				575						
Tyr	His	Lys	Met	Glu	Met	Lys	Gln	Ala	Ile	Glu	Leu	Val	Glu	Ser	Gly
580					585				590						
Gly	Met	Gly	Lys	Ile	Pro	Ser	Ala	Val	Ser	Thr	Val	Ser	Met	Gln	Asn
595					600				605						
Ile	His	Pro	Lys	Ser	Leu	Pro	Ser	Glu	Arg	Ile	Leu	Pro	Ala	Leu	Ser
610					615				620						
Lys	Asp	Lys	Glu	Glu	Glu	Ile	Arg	Lys	Ile	Leu	Arg	Asn	Asn	Leu	Gln
625					630				635				640		
Lys	Thr	Arg	Gln	Arg	Leu	Arg	Ser	Tyr	Asn	Arg	His	Thr	Leu	Val	Ala
645					650				655						
Asp	Pro	Tyr	Glu	Glu	Ala	Trp	Asn	Gln	Met	Leu	Leu	Arg	Arg	Gln	Lys
660					665				670						
Ala	Arg	Gln	Leu	Glu	Gln	Lys	Ile	Asn	Asn	Tyr	Leu	Thr	Val	Pro	Ala
675					680				685						
His	Lys	Leu	Asp	Ser	Pro	Thr	Met	Ser	Arg	Ala	Arg	Ile	Gly	Ser	Asp
690					695				700						
Pro	Leu	Ala	Tyr	Glu	Pro	Lys	Glu	Asp	Leu	Pro	Val	Ile	Thr	Ile	Asp
705					710				715				720		
Pro	Ala	Ser	Pro	Gln	Ser	Pro	Glu	Ser	Val	Asp	Leu	Val	Asn	Glu	Glu
725					730				735						
Leu	Lys	Gly	Lys	Val	Leu	Gly	Leu	Ser	Arg	Asp	Pro	Ala	Lys	Val	Ala
740					745				750						
Glu	Glu	Asp	Glu	Asp	Asp	Asp	Gly	Gly	Ile	Met	Met	Arg	Ser	Lys	Glu
755					760				765						
Thr	Ser	Ser	Pro	Gly	Thr	Asp	Asp	Val	Phe	Thr	Pro	Ala	Pro	Ser	Asp
770					775				780						
Ser	Pro	Ser	Ser	Gln	Arg	Ile	Gln	Arg	Cys	Leu	Ser	Asp	Pro	Gly	Pro
785					790				795				800		
His	Pro	Glu	Pro	Gly	Glu	Glu	Pro	Phe	Phe	Pro	Lys	Gly	Gln		
805					810				815						

<210> SEQ ID NO 51

<211> LENGTH: 4606

<212> TYPE: DNA

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 51

atcagtccgc	taccgcggga	ccccgggctg	tgtcaggctc	ggaagtggag	ctggagctga	60
ccgaaggagc	tgcgcgtggg	ccgggggtgga	acacagctc	gcaccgggg	cttaaggagg	120
ggtgtgttgc	tgcgtggctca	ctgcgagccg	cctccagcct	gcgccggggc	acttcccgc	180
caggccgtct	gcctctgtct	tccgcctccc	tccatttccc	cggaatctca	gcccggcg	240
gctggaccct	ggtcttcctg	tgggtgggaa	agccccagcc	tttttcctgc	tataactcctt	300
tatccccagcc	tggcccttcca	gccttctctc	gccttccct	gcggagtggg	tctcacccctt	360
tcccggtctt	agccccaggct	ccttaccct	cccccttagc	cgagttccag	caccctgggt	420

-continued

attttcctta gccaaaggc cstatccttct tccctccccc ccttgeccgt ggtccagcct	480
atccggccct tccccctctct tcaggacgca gtttccccga ttcccttctc tcagectgaa	540
gttcttagatc ctcccggtt cttcccttagt ccctccaagg tacttaaccc cttttgcag	600
tttggacttt gagattgtga atctctagaa aacgaaaata cttctctaaa acatgaactt	660
ttccttagaga ttgccccagt ctcaccccat ctgcaggact ccttgcctacc tatgtccccag	720
gtcttacaeg caattcgcta aaggctgagg cctccttaggc tggggatca gtatgtatgc	780
tccgtggtcc ggctgtctggg gatttcatcc acctcggttc ttccctctc tgctgggt	840
gtttgccttg gtgggactgc tacctgttct caggagccac ggcctccagc acagccctac	900
tgccagcacc atcagagggtt ctgaaccacc cgggaaacgc tcaattgggg atgtcaccac	960
ggcaccatca gagcctctcc atcgcccaga tgaccacaat ttgaccaact taatcattga	1020
acatggcggt aagccatctc ggaaggcctt ccacgtccgt gacattgact accctcacgt	1080
gcgeacaccc tttgagatct cccctctggat tccctctggcc tgcctcatga agataggttt	1140
ccatgtgatc cccaccatct caagcatcg cccggagagc tgcctgctga tcgtatgtgg	1200
gctgctggtg gggggcctga tcaaggcggt cgagagacg ccccccctcc tgcaatcaga	1260
cgtttcttc ctttccctgc tgccacccat catcctggat gcaggctact tcctgcctct	1320
gcggcagtcc acggagaacc tgggacccat cctgatctt gctgtggtgg gcacactgtg	1380
gaatgcgttc ttccctgggtg gctccctgtt cggcgatgc ctggggcg gcgagcagat	1440
caacaacatt ggcctgctgg acaccctgtc cttggcagc atcatctctg ccgtggaccc	1500
tgtggctgtg ctggccgtct tcgaggagat ccacatcaac gaactgtac acatcttgc	1560
cttcggggag tccctgctca acgtgccgt cactgtggtc ctctatcacc tctttgagga	1620
atttgccagc tatgactctg tgggacccat ggacatcttc ctgcgttcc tgagttctt	1680
cgtggatggcc ctccgggggg tggttgggg cgtggatctac ggggtatcg cagcttcac	1740
ctcccgattt acctctcaca tccgggtcat cgagccgctc ttccgttcc tctacagcta	1800
catggcttac ctgtcagctg agcttccca cctgtccgga atcatggccc tcacgcctc	1860
aggagtagtg atgcgcccct atgtggaggc caacatctcc cacaaatccc acaccacca	1920
caagtacttc ctgaagatgt ggagcagcgt cagttagacc ctcatcttca tcttcctgg	1980
cgtctccact gtggccgggt cccaccatgt gaactggacc ttccgtatca gcaccctgt	2040
cttctgcctc atcgccccgg tgctgggtgt cctggatctc acctgggtca tcaacaagtt	2100
ccgcattgtc aagctgaccc ccaaggacca gttcatcatt gcctatgggg gcctgcgagg	2160
ggccatcgcc ttctcgctgg gctacccctt ggacaagaag cacttccca tggatgtac	2220
gttccctacc gccatcatca ccgtcatctt ttccacccgtc ttgtgcagg gcatgaccat	2280
tccggccctg gtccgacctgt tggccgtgaa gaaaaagcaa gaaacaaagc gctccatcaa	2340
cgaggagatc cacacacgt tccctggacca cttctgtaca ggcacatcgagg acatctgtgg	2400
tcattatggc caccatcaact ggaaggacaa gtcacaccgt tttataaaga agtatgtgaa	2460
gaagtgcctg atagcaggag agcgctccaa agagccccag ctcattgcct tctaccacaa	2520
gatggagatg aacgaggcca ttgagctggt ggagagcggg ggcacatggca agatcccatc	2580
tgctgtctcc actgtctcta tgcagaacat ccaccccaag gctgtgactt cagacccat	2640
cttgcacca ctgtccaaagg acaaggagga agagatccgc aaaatcttga ggagcaatct	2700

-continued

cgagaaaacc cggcaacggc tgccgttccta taacagacac acgctgggtgg ccgaccccta 2760
cgaggaaagcc tggAACCCAGA tgctgtcccg gaggcagaag gccccggcagc tggagcagaa 2820
gatcaccaac tacccgtacag taccggccca caagctggac tcgcccaccc tgtctcgccc 2880
ccgcataggc tcagacccac tggcctatga acccaaggca gacctgcctg tcacaccat 2940
tgaccggcc tccccacagt cccctgagtc tgtggacctg gtgaacgagg agttgaaggg 3000
caagggtccta gggcttaaaccc ggggtcccaag ggtgactccg gaggaggagg aggaagatga 3060
ggacggcatc atcatgattt ggagcaagga gcctctatcc ccaggcaccg acgacgttt 3120
ccccctgg a tcaagtgaca gcggcagtc ccagagaata caacgctgcc tcagtgcacc 3180
aggccccac cctgagcctg gggagggaga gccttcatc cccaaaggac agtagagcca 3240
ggctagtggg tggccacag actctccacag gagagtgggg gctggggcgct cctcttgtct 3300
cctgacctgg attggctctg cccctttctt ccgcattggca gctggggccca cagccccac 3360
tgcagcacgg ctggccctg cctcctggga agggcttctt cctacagaac tgccttccca 3420
atccatccat ttggcggacg gctggctgc tgaggcaggc cccgcggcat ctctagatcc 3480
aggctctccc atagtggAAC cagggtctctt caaggccctc ctcccttcca gcccctcc 3540
atccagacca actctccatc ttgtatcatc atttctaacc aaagagccac tggctccact 3600
gtggtcccaag caggaaagga gggagctgag cttccgtgg ccgtgactgg gcccctgtat 3660
cagaggacag caggcctgtg gctggaaac agaccctgtg cccatagcag agctggccac 3720
ggcagcgcca cgggttccccc gtagcctcg gtgtctggc caccacccga gctgtccctta 3780
ctcagagctg cgggctgagg ggtctctgat gttctctgc tgggctggg ggcacatcc 3840
acagtgcacat gattgggggtg atgaggacca ctggcccta gtggggctta cacgggaggg 3900
actgtcccgag ggcgtggaaa gtgtccctta cttacacttg gtttatttgc acctcaactgt 3960
aagtatgatt ctgtttgcac agaggactc ttgtttctg accctttgag tccccatgc 4020
cagtcagacg tctgcctt cccctcgat cttccctgtg tctaaacaccg tcccttaaat 4080
ggaaacctctg ctgtgggggg cccttgcaca cagcaccatg tctgtatgtgg cccaaacctgt 4140
catgactcaa gtcaccacggc agggggccctc ctggtaacctc cttgtctttt gggcctgttt 4200
ttactctgt tggcccgatcc taccaccacc ctttatctgg accactctct ttcctgtcc 4260
ctgttagagta gaaagtactg gaccccttag aggcccttgc tcacaggcac acactcccta 4320
ccggccacat ggctggggcc aagcccatgg gtatctggcc aaccctggg cactccctcc 4380
ccaggatcgat gcaataatag tcagagggtc cttccctggg gatggggccca tgggtgcatt 4440
ggccagctgt ctccctcttca tgcaagtgtct gttttggca ggagttggca ggcagggtga 4500
catcgacaaac catgtacaaa gccaccgcag ctgtgcgcgc tctgcgcct gtacaaagga 4560
aactgaacct ttttcatatt ctaataatc aatgtgagtt ttgtat 4606

<210> SEQ ID NO 52

<211> LENGTH: 820

<212> TYPE: PRT

<212> TYPE: TRI
<213> ORGANISM: *Mus musculus*

<400> SEQUENCE: 52

```

Met Met Leu Arg Trp Ser Gly Val Trp Gly Phe His Pro Pro Arg Ile
          5                   10                  15

```

Phe Pro Ser Leu Leu Val Val Val Ala Leu Val Gly Leu Leu Pro Val

-continued

20	25	30
Leu Arg Ser His Gly Leu Gln His Ser Pro Thr Ala Ser Thr Ile Arg		
35	40	45
Gly Ser Glu Pro Pro Arg Glu Arg Ser Ile Gly Asp Val Thr Thr Ala		
50	55	60
Pro Ser Glu Pro Leu His Arg Pro Asp Asp His Asn Leu Thr Asn Leu		
65	70	75
Ile Ile Glu His Gly Gly Lys Pro Ser Arg Lys Ala Phe Pro Val Leu		
85	90	95
Asp Ile Asp Tyr Pro His Val Arg Thr Pro Phe Glu Ile Ser Leu Trp		
100	105	110
Ile Leu Leu Ala Cys Leu Met Lys Ile Gly Phe His Val Ile Pro Thr		
115	120	125
Ile Ser Ser Ile Val Pro Glu Ser Cys Leu Leu Ile Val Val Gly Leu		
130	135	140
Leu Val Gly Gly Leu Ile Lys Gly Val Gly Glu Thr Pro Pro Phe Leu		
145	150	155
Gln Ser Asp Val Phe Phe Leu Phe Leu Leu Pro Pro Ile Ile Leu Asp		
165	170	175
Ala Gly Tyr Phe Leu Pro Leu Arg Gln Phe Thr Glu Asn Leu Gly Thr		
180	185	190
Ile Leu Ile Phe Ala Val Val Gly Thr Leu Trp Asn Ala Phe Phe Leu		
195	200	205
Gly Gly Leu Leu Tyr Ala Val Cys Leu Val Gly Gly Glu Gln Ile Asn		
210	215	220
Asn Ile Gly Leu Leu Asp Thr Leu Leu Phe Gly Ser Ile Ile Ser Ala		
225	230	235
Val Asp Pro Val Ala Val Leu Ala Val Phe Glu Glu Ile His Ile Asn		
245	250	255
Glu Leu Leu His Ile Leu Val Phe Gly Glu Ser Leu Leu Asn Asp Ala		
260	265	270
Val Thr Val Val Leu Tyr His Leu Phe Glu Glu Phe Ala Ser Tyr Asp		
275	280	285
Ser Val Gly Ile Ser Asp Ile Phe Leu Gly Phe Leu Ser Phe Phe Val		
290	295	300
Val Ala Leu Gly Gly Val Phe Val Gly Val Val Tyr Gly Val Ile Ala		
305	310	315
Ala Phe Thr Ser Arg Phe Thr Ser His Ile Arg Val Ile Glu Pro Leu		
325	330	335
Phe Val Phe Leu Tyr Ser Tyr Met Ala Tyr Leu Ser Ala Glu Leu Phe		
340	345	350
His Leu Ser Gly Ile Met Ala Leu Ile Ala Ser Gly Val Val Met Arg		
355	360	365
Pro Tyr Val Glu Ala Asn Ile Ser His Lys Ser His Thr Thr Ile Lys		
370	375	380
Tyr Phe Leu Lys Met Trp Ser Ser Val Ser Glu Thr Leu Ile Phe Ile		
385	390	395
Phe Leu Gly Val Ser Thr Val Ala Gly Ser His Gln Trp Asn Trp Thr		
405	410	415
Phe Val Ile Ser Thr Leu Leu Phe Cys Leu Ile Ala Arg Val Leu Gly		
420	425	430

-continued

Val Leu Val Leu Thr Trp Phe Ile Asn Lys Phe Arg Ile Val Lys Leu
 435 440 445
 Thr Pro Lys Asp Gln Phe Ile Ile Ala Tyr Gly Gly Leu Arg Gly Ala
 450 455 460
 Ile Ala Phe Ser Leu Gly Tyr Leu Leu Asp Lys Lys His Phe Pro Met
 465 470 475 480
 Cys Asp Leu Phe Leu Thr Ala Ile Ile Thr Val Ile Phe Phe Thr Val
 485 490 495
 Phe Val Gln Gly Met Thr Ile Arg Pro Leu Val Asp Leu Leu Ala Val
 500 505 510
 Lys Lys Lys Gln Glu Thr Lys Arg Ser Ile Asn Glu Glu Ile His Thr
 515 520 525
 Gln Phe Leu Asp His Leu Leu Thr Gly Ile Glu Asp Ile Cys Gly His
 530 535 540
 Tyr Gly His His His Trp Lys Asp Lys Leu Asn Arg Phe Asn Lys Lys
 545 550 555 560
 Tyr Val Lys Lys Cys Leu Ile Ala Gly Glu Arg Ser Lys Glu Pro Gln
 565 570 575
 Leu Ile Ala Phe Tyr His Lys Met Glu Met Lys Gln Ala Ile Glu Leu
 580 585 590
 Val Glu Ser Gly Gly Met Gly Lys Ile Pro Ser Ala Val Ser Thr Val
 595 600 605
 Ser Met Gln Asn Ile His Pro Lys Ala Val Thr Ser Asp Arg Ile Leu
 610 615 620
 Pro Ala Leu Ser Lys Asp Lys Glu Glu Glu Ile Arg Lys Ile Leu Arg
 625 630 635 640
 Ser Asn Leu Gln Lys Thr Arg Gln Arg Leu Arg Ser Tyr Asn Arg His
 645 650 655
 Thr Leu Val Ala Asp Pro Tyr Glu Glu Ala Trp Asn Gln Met Leu Leu
 660 665 670
 Arg Arg Gln Lys Ala Arg Gln Leu Glu Gln Lys Ile Thr Asn Tyr Leu
 675 680 685
 Thr Val Pro Ala His Lys Leu Asp Ser Pro Thr Leu Ser Arg Ala Arg
 690 695 700
 Ile Gly Ser Asp Pro Leu Ala Tyr Glu Pro Lys Ala Asp Leu Pro Val
 705 710 715 720
 Ile Thr Ile Asp Pro Ala Ser Pro Gln Ser Pro Glu Ser Val Asp Leu
 725 730 735
 Val Asn Glu Glu Leu Lys Gly Lys Val Leu Gly Leu Asn Arg Gly Pro
 740 745 750
 Arg Val Thr Pro Glu Glu Glu Glu Asp Glu Asp Gly Ile Ile Met
 755 760 765
 Ile Arg Ser Lys Glu Pro Ser Ser Pro Gly Thr Asp Asp Val Phe Thr
 770 775 780
 Pro Gly Ser Ser Asp Ser Pro Ser Ser Gln Arg Ile Gln Arg Cys Leu
 785 790 795 800
 Ser Asp Pro Gly Pro His Pro Glu Pro Gly Glu Gly Glu Pro Phe Ile
 805 810 815
 Pro Lys Gly Gln
 820

We claim:

1. A method of promoting regeneration, comprising
 - (a) providing a population of naturally regenerating cells;
 - (b) determining the membrane potential or pH range permissive for regeneration in said population of regenerating cells;
 - (c) providing a population of non-regenerating cells;
 - (d) determining the membrane potential or pH range of said population of non-regenerating cells; and
 - (e) contacting the population of non-regenerating cells with an agent that modulates ion flux mediated by a class of ion transporter proteins, which agent modifies the membrane potential or pH of one or more cells in the population of non-regenerating cells to the range permissive for regeneration as determined in step (b); thereby promoting regeneration of one or more cells in the population of cells.
2. A method of promoting regeneration, comprising
 - (a) determining the membrane potential or pH range permissive for regeneration in a population of regenerating cells;
 - (b) providing a population of non-regenerating cells;
 - (c) determining the membrane potential or pH range of said population of non-regenerating cells; and
 - (d) contacting the population of non-regenerating cells with an agent that modulates ion flux mediated by a class of ion transporter proteins, which agent modifies the membrane potential or pH of one or more cells in the population of non-regenerating cells to the range permissive for regeneration as determined in step (b); thereby promoting regeneration of one or more cells in the population of cells.
3. The method of claim 1, wherein step (b) comprises
 - (i) providing a population of regenerating cells;
 - (ii) contacting said population of cells with an agent which is a voltage sensitive agent that produces a detectable signal; and
 - (iii) measuring the detectable signal to calculate an average membrane potential or pH of said population of cells during regeneration.
4. The method of claim 1, wherein steps (a) and (b) are repeated for multiple types of naturally regenerating cells and the average membrane potential or pH range are used in step (e).
5. The method of claim 1, wherein step (d) comprises
 - (i) providing a population of non-regenerating cells;
 - (ii) contacting said population of cells with an agent which is a voltage sensitive agent that produces a detectable signal; and
 - (iii) measuring the detectable signal to calculate an average membrane potential or pH of said population of cells.
6. The method of claim 1, wherein the method is an in vitro method and the population of non-regenerating cells are in culture.
7. The method of claim 1, wherein the method is an in vitro method and the population of non-regenerating cells are in a preparation of tissue in culture.
8. The method of claim 1, wherein the agent inhibits ion flux mediated by the class of transporter proteins.
9. The method of claim 1, wherein the agent promotes ion flux mediated by the class of transporter proteins.
10. The method of claim 1, wherein the agent is an ion channel protein or a nucleotide construct that encodes an ion channel protein.

11. The method of claim 10, wherein the ion transporter protein is a hyperpolarizing transporter.
12. The method of claim 10, wherein the ion transporter protein is a depolarizing transporter.
13. The method of claim 10, wherein the ion transporter protein is an H⁺ pump.
14. The method of claim 13, wherein the H⁺ pump is a V-ATPase H⁺ pump.
15. A method of promoting regeneration, comprising
 - (a) providing a population of cells of known membrane potential or pH; and
 - (b) contacting the population of cells with an agent that modulates ion flux mediated by a class of ion transporter proteins, which agent modifies the membrane potential or pH of one or more cells in the population of cells to a range permissive for regeneration; thereby promoting regeneration of one or more cells in the population of cells.
16. The method of claim 15, wherein contacting the population of cells with the agent promotes dedifferentiation.
17. The method of claim 16, wherein the dedifferentiated cells are cultured for a time sufficient to allow proliferation.
18. The method of claim 15, wherein the membrane potential range permissive for regeneration is between -70 mV and 30 mV.
19. A method of promoting regeneration, comprising
 - (a) providing a population of cells; and
 - (b) contacting the population of cells with an agent that increases ion flux mediated by a V-ATPase H⁺ pump, which agent promotes relative hyperpolarization of cell membranes of one or more cells in the population of cells, thereby promoting regeneration of one or more cells in the population of cells.
20. The method of claim 19, wherein contacting the population of cells with the agent promotes dedifferentiation.
21. The method of claim 20, wherein the dedifferentiated cells are cultured for a time sufficient to allow proliferation.
22. The method of claim 19, wherein the agent is a nucleotide construct encoding the V-ATPase H⁺ pump.
23. A method of promoting dedifferentiation, comprising
 - (a) determining the membrane potential or pH range permissive for dedifferentiation in a population of dedifferentiated cells;
 - (b) providing a population of differentiated cells;
 - (c) determining the membrane potential or pH range of said population of differentiated cells; and
 - (d) contacting the population of differentiated cells with an agent that modulates ion flux mediated by a class of ion transporter proteins, which agent modifies the membrane potential or pH of one or more cells in the population of differentiated cells to the range permissive for dedifferentiation as determined in step (b); thereby promoting dedifferentiation of one or more cells in the population of cells.
24. A method of inhibiting dedifferentiation, comprising
 - (a) determining the membrane potential or pH range permissive for dedifferentiation in a population of dedifferentiated cells;
 - (b) providing a population of cells;
 - (c) determining the membrane potential or pH range of said population of cells; and
 - (d) contacting the population of cells with an agent that modulates ion flux mediated by a class of ion transporter proteins, which agent modifies the membrane potential

or pH of one or more cells in the second population of cells out of the range permissive for dedifferentiation as determined in step (b); thereby inhibiting dedifferentiation of one or more cells in the population of cells.

25. The method of claim **23**, further comprising culturing the population of cells, which population of cells includes one or more dedifferentiated cells, wherein said culturing promotes regeneration.

26. The method of claim **23**, wherein the compound inhibits ion flux mediated by the class of transporter proteins.

27. The method of claim **23**, wherein the compound promotes ion flux mediated by the class of transporter proteins.

28. A method of identifying progenitor cells, comprising
 (a) contacting a population of cells with an agent, which agent is a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane;
 (b) identifying, in the population of cells, one or more cells having a membrane potential of greater than or equal to -20 mV,

wherein the one or more cells having a membrane potential of greater than or equal to -20 mV are identified as candidate progenitor cells.

29. A method of identifying progenitor cells, comprising
 (a) contacting a population of cells with an agent, which agent is a pH sensitive agent that produces a detectable signal;
 (b) identifying, in the population of cells, one or more cells which have an intracellular pH of less than or equal to 6.7,

wherein the one or more cells having an intracellular pH of less than or equal to 6.7 are identified as candidate progenitor cells.

30. The method of claim **28**, wherein the method is an in vitro method and the population of cells is in culture.

31. The method of claim **28**, wherein the method is an in vivo method and the cells are in an animal.

32. The method of claim **28**, wherein the candidate progenitor cells have a membrane potential greater than or equal to -20 mV and less than or equal to 30 mV.

33. The method of claim **28**, further comprising
 contacting the population of cells with a pH sensitive agent that produces a detectable signal; and
 identifying, in the population of cells, one or more cells with an intracellular pH of less than or equal to 6.7,
 wherein the one or more cells having both a membrane potential of greater than or equal to -20 mV and an intracellular pH of less than or equal to 6.7 are candidate progenitor cells.

34. The method of claim **28**, further comprising
 (c) separating candidate progenitor cells from the population of the cells.

35. The method of claim **28**, further comprising
 (d) culturing candidate progenitor cells separated in (c) to produce a population of cells enriched for candidate progenitor cells.

36. The method of claim **28**, wherein the detectable signal is a fluorescent signal.

37. A method of identifying progenitor cells, comprising
 (a) contacting a population of cells with a first agent, which first agent is a voltage sensitive agent that produces a detectable signal in response to a depolarized cell membrane;
 (b) contacting the population of cells with a second agent, which second agent is a pH sensitive agent that produces a detectable signal;

(c) identifying, in the population of cells, one or more cells having a membrane potential of greater than or equal to -20 mV, and an intracellular pH of less than or equal to 6.7,

wherein the one or more cells having both a membrane potential of greater than or equal to -20 mV and an intracellular pH of less than or equal to 6.7 are identified as candidate progenitor cells.

38. The method of claim **37**, wherein the population of cells is contacted simultaneously with the first agent and the second agent.

39. The method of claim **37**, wherein the population of cells is contacted sequentially with the first agent and the second agent.

40. The method of claim **37**, further comprising
 (d) separating candidate progenitor cells from the population of the cells.

41. The method of claim **40**, further comprising
 (e) culturing candidate progenitor cells separated in (d) to produce a population of cells enriched for candidate progenitor cells.

42. A method of separating progenitor cells from an animal or tissue, comprising

(a) contacting an animal or tissue with an agent, which agent is a voltage sensitive agent that produces a detectable signal in cells in response to a depolarized cell membrane;

(b) identifying, in the animal or tissue, one or more cells having a membrane potential of greater than or equal to -20 mV,

wherein the one or more cells having a membrane potential of greater than or equal to -20 mV are identified as candidate progenitor cells;

(c) removing identified candidate progenitor cells, thereby separating candidate progenitor cells from the animal or tissue.

43. A method of separating progenitor cells from an animal or tissue, comprising

(a) contacting an animal or tissue with an agent, which agent is a pH sensitive agent that produces a detectable signal;

(b) identifying, in the animal or tissue, one or more cells with an intracellular pH of less than or equal to 6.7,

wherein the one or more cells having an intracellular pH of less than or equal to 6.7 are identified as candidate progenitor cells;

(c) removing identified candidate progenitor cells, thereby separating candidate progenitor cells from the animal or tissue.

44. The method of claim **42**, further comprising
 contacting the animal or tissue with an agent, which agent is a pH sensitive agent that produces a detectable signal; and
 identifying, in the animal or tissue, one or more cells with an intracellular pH of less than or equal to 6.7,

wherein the one or more cells having both a membrane potential of greater than or equal to -20 mV and an intracellular pH of less than or equal to 6.7 are identified as candidate progenitor cells.

45. The method of claim **42**, wherein removing identified candidate progenitor cells comprises dissecting out candidate progenitor cells, thereby removing candidate progenitor cells from the animal or tissue.

46. The method of claim **42**, wherein removing identified candidate progenitor cells comprises dissociating the animal or tissue; and

sorting candidate progenitor cells, thereby separating candidate progenitor cells from the animal or tissue.

47. The method of claim **46**, wherein sorting the candidate progenitor cells comprises an automated method of sorting candidate progenitor cells based on the detectable signal.

48. A method for identifying a candidate class of ion transporter proteins which mediate ion flux during a particular biological process, comprising

(a) providing a population of cells for measuring a particular biological process;

(b) contacting the population of cells with a first compound that modulates ion flux mediated by a first class of ion transporter proteins;

(c) measuring the particular biological process in the population of cells in the presence of the compound versus the absence of the compound;

(d) determining whether the compound which modulates ion flux mediated by the first class of ion transporter proteins changes the particular biological process in the population of cells, thereby identifying a candidate class of ion transporters which may mediate ion flux during the particular biological process;

(e) providing a second population of cells for measuring said particular biological process;

(f) contacting the second population of cells with a second compound that modulates ion flux mediated by a second class of ion transporter proteins, which second class of ion transporter proteins comprises a subset of the first class of ion transporter proteins;

(g) measuring the particular biological process in the population of cells in the presence of the second compound versus the absence of the second compound; and

(h) determining whether the second compound which modulates ion flux mediated by the second class of ion transporter proteins changes the particular biological process in the population of cells,

thereby identifying a class of ion transporters which mediate ion flux during the particular biological process.

49. The method of claim **48**, wherein the compound inhibits ion flux mediated by the class of transporter proteins.

50. The method of claim **48**, wherein the compound promotes ion flux mediated by the class of transporter proteins.

51. The method of claim **48**, wherein the particular biological process is selected from cell proliferation, cell differentiation, apoptosis, cell survival, cell migration, regeneration, or dedifferentiation.

52. The method of claim **48**, wherein measuring the particular biological process comprises measuring a change in gene expression, a change in protein expression, or a change in morphology.

53. The method of claim **48**, wherein both the first compound and the second compound change the particular biological process, thereby identifying a candidate class of ion transporter proteins which mediate ion flux during the particular biological process, and which candidate class of ion transporter proteins is a subset of the first class of ion transporter proteins.

54. The method of claim **48**, furthering comprising providing a population of cells for measuring said particu-

lar biological process;
inhibiting expression or activity of an ion transporter protein, which ion transporter protein is a member of the candidate class of ion transporter proteins which mediate the particular biological process;

measuring the particular biological process in the popula-

tion of cells; and
determining whether inhibition of the expression or activi-

ty of said ion transporter protein changes the particular biological process in the population of cells,

thereby identifying an ion transporter protein which mediates ion flux during the particular biological process.

55. The method of claim **54**, wherein inhibiting the expres-

sion or activity of the ion transporter protein comprises
contacting the population of cells with an agent that spe-

cifically inhibits the expression or activity of the ion

transporter protein and does not substantially inhibit the

expression or activity of other ion transporter proteins

that are a member of the candidate class of ion trans-

porter proteins.

56. A method for modulating a particular biological pro-

cess, comprising

(a) identifying a candidate class of ion transporter proteins
which may mediate ion flux during a particular biologi-

cal process according to the method of claim **48**;

(b) contacting a population of cells with a compound that
modulates the expression or activity of one or more ion
transporter proteins of the candidate class of ion trans-

porter proteins; thereby modulating the particular bio-

* * * * *