

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
24 April 2003 (24.04.2003)

PCT

(10) International Publication Number
WO 03/033646 A2

(51) International Patent Classification⁷: **C12N**
(21) International Application Number: PCT/US02/24589
(22) International Filing Date: 31 July 2002 (31.07.2002)
(25) Filing Language: English
(26) Publication Language: English

(30) Priority Data:
60/308,958 31 July 2001 (31.07.2001) US
60/345,846 9 November 2001 (09.11.2001) US

(71) Applicant (for all designated States except US): **PROTEOLOGICS, INC.** [US/US]; 40 Ramland Road South, Orangeburg, NY 10962-2614 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GREENER, Tsvika** [IL/IL]; Nes-Ziona (IL). **MOSKOWITZ, Haim** [IL/IL]; Nes-Ziona (IL). **REISS, Yuval** [IL/IL]; Kiron (IL). **ALROY, Iris** [IL/IL]; Rehovot (IL).

(74) Agent: **VARMA, Anita**; Ropes & Gray, One International Place, Boston, MA 02110-2624 (US).

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

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

Published:

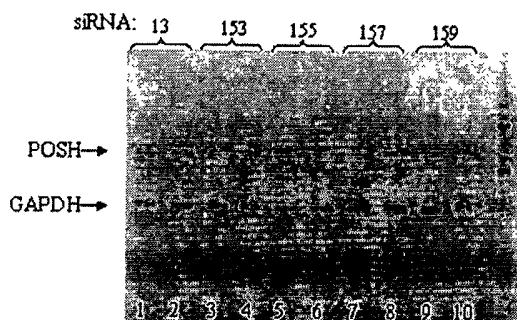
— without international search report and to be republished upon receipt of that report

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

(54) Title: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL MATURATION

Reduction in Full Length PRT3 mRNA by siRNA Duplexes

(57) Abstract: This application describes a family of nucleic acid sequences and proteins encoded thereby that play a role in viral maturation: the Alternate Viral Maturation Scaffolding Protein, or the AVMSP family of proteins.



WO 03/033646 A2

COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL MATURATION

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Nos. 60/308,958, filed July 31, 2001, and 60/345,846, filed November 9, 2001. The entire contents of these applications are herein incorporated by reference.

BACKGROUND

Viral maturation requires the proteolytic processing of viral proteins, such as Gag, and the activity of the host proteins. It is believed that cellular machineries for exo/endocytosis and for ubiquitin conjugation may be involved in the maturation. In particular, the assembly and subsequent budding of retroviruses, rhabdoviruses, and filoviruses depends on the Gag polyprotein. After its synthesis, Gag is targeted to the plasma membrane where it induces budding of nascent virus particles.

The role of ubiquitin in virus assembly was suggested by Dunigan et al. (1988, *Virology* 165, 310, Meyers et al. 1991, *Virology* 180, 602), who observed that mature virus particles were enriched in unconjugated ubiquitin. More recently, it was shown that proteasome inhibitors suppress the release of HIV-1, HIV-2 and virus-like particles derived from SIV and RSV Gag. Also, inhibitors affect Gag processing and maturation into infectious particles (Schubert et al 2000, *PNAS* 97, 13057, Harty et al. 2000, *PNAS* 97, 13871, Strack et al. 2000, *PNAS* 97, 13063, Patnaik et al. 2000, *PNAS* 97, 13069).

It is well known in the art that ubiquitin-mediated proteolysis is the major pathway for the selective, controlled degradation of intracellular proteins in eukaryotic cells. Ubiquitin modification of a variety of protein targets within the cell appears to be important in a number of basic cellular functions such as regulation of gene expression, regulation of the cell-cycle, modification of cell surface receptors, biogenesis of ribosomes, and DNA repair. One major function of the ubiquitin-mediated system is to control the half-lives of cellular proteins. The half-life of different proteins can range from a few minutes to several days, and can vary considerably depending on the cell-type, nutritional and environmental conditions, as well as the stage of the cell-cycle.

Targeted proteins undergoing selective degradation, presumably through the actions of a ubiquitin-dependent proteasome, are covalently tagged with ubiquitin through the formation of an isopeptide bond between the C-terminal glycyl residue of ubiquitin and a

specific lysyl residue in the substrate protein. This process is catalyzed by a ubiquitin-activating enzyme (E1) and a ubiquitin-conjugating enzyme (E2), and in some instances may also require auxiliary substrate recognition proteins (E3s). Following the linkage of the first ubiquitin chain, additional molecules of ubiquitin may be attached to lysine side chains of the previously conjugated moiety to form branched multi-ubiquitin chains.

The conjugation of ubiquitin to protein substrates is a multi-step process. In an initial ATP requiring step, a thioester is formed between the C-terminus of ubiquitin and an internal cysteine residue of an E1 enzyme. Activated ubiquitin is then transferred to a specific cysteine on one of several E2 enzymes. Finally, these E2 enzymes donate ubiquitin to protein substrates. Substrates are recognized either directly by ubiquitin-conjugated enzymes or by associated substrate recognition proteins, the E3 proteins, also known as ubiquitin ligases.

SUMMARY

It is proposed that a variety of proteins, including ubiquitin protein ligases and proteins involved in membrane trafficking, are recruited for the process of viral maturation (including, for example, assembly, budding and release) by direct or indirect interaction with viral proteins, for example Gag proteins. The ligase then ubiquitinates viral and/or cellular proteins that are part of the membrane remodeling machinery. For example, a number of Gag protein motifs such as PxxP, PxxY, PPxY and YxxL, are known to recruit proteins involved in viral maturation.

To this end, in certain embodiments, the invention provides a family of nucleic acid sequences and proteins encoded thereby that play a role in viral maturation: the Alternate Viral Maturation Scaffolding Protein, or the AVMSPP family of proteins. Broadly, AVMSPP polypeptides comprise a first domain or functional role and a second domain. The first domain or functional role is selected from the following: SH2, SH3, or membrane spanning ("membrane"), or functions as a receptor. A preferred AVMSPP also comprises a second domain that is a RING domain. Accordingly, different categories of AVMSPPs may be referred to as RING-SH3 proteins and nucleic acids, RING-SH2 proteins and nucleic acids, RING-membrane proteins and nucleic acids and RING-receptor polypeptides and nucleic acids. The first domain and second domain may be found in any order within the AVMSPP sequence (i.e. the first domain need not be N-terminal to the second domain). It is understood that polypeptides that function as receptors will often have membrane or other domains. In certain embodiments AVMSPP proteins comprise a C2 domain.

In further aspects, in cells infected with viruses that utilize a Gag-dependent pathway for assembly, budding and/or release, AVMSPs, act to assemble complexes of proteins that mediate release. AVMSP complexes may, for example, stimulate ubiquitylation of certain proteins, stimulate membrane fusion, stimulate assembly of viral particles, or a combination of the preceding. As one of skill in the art can readily appreciate, any single AVMSP may form multiple different complexes at different times.

In additional aspects, the invention provides nucleic acid sequences and proteins encoded thereby, as well as probes derived from the nucleic acid sequences, antibodies directed to the encoded proteins, diagnostic methods for detecting cells infected with a virus, and assays for identifying agents having an antiviral activity.

In one aspect, the invention provides a RING-SH3 nucleic acid, such as an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence encoding a RING-SH3 protein, such as a sequence of SEQ ID Nos: 1-3, or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos. 40-99 or a sequence complementary thereto or up to the full length of the gene of which said sequence is a fragment. In a further embodiment, the RING-SH3 nucleic acid comprises a nucleic acid encoding an amino acid sequence as set forth in SEQ ID Nos. 1-39 or a nucleic acid complement thereof. In a related embodiment, the encoded amino acid sequence is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive amino acids up to the full length of one of SEQ ID Nos: 1-39. In yet another embodiment, the RING-SH3 nucleic acid is an isolated nucleic acid encoding a polypeptide comprising a RING domain and an SH3 domain. In a preferred embodiment, the RING-SH3 nucleic acid is a PRT3 nucleic acid of SEQ ID NOs:40-44 or a functional variant thereof.

In a further aspect, the invention provides a RING-SH2 nucleic acid, such as an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence encoding a RING-SH2 protein, such as a sequence of SEQ ID Nos: 45-46, or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos.45-46, or a sequence complementary thereto or up

to the full length of the gene of which said sequence is a fragment. In a further embodiment, the RING-SH2 nucleic acid comprises a nucleic acid encoding an amino acid sequence as set forth in SEQ ID Nos. 4-5, or a nucleic acid complement thereof. In a related embodiment, the encoded amino acid sequence is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive amino acids up to the full length of one of SEQ ID NOS: 4-5. In yet another embodiment, the RING-SH2 nucleic acid is an isolated nucleic acid encoding a polypeptide comprising a RING domain and an SH2 domain.

In a further aspect, the invention provides a RING-membrane nucleic acid, such as an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence encoding a RING-membrane protein, such as a sequence of SEQ ID Nos: 47-56, or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos.47-56, or a sequence complementary thereto or up to the full length of the gene of which said sequence is a fragment. In a further embodiment, the RING-membrane nucleic acid comprises a nucleic acid encoding an amino acid sequence as set forth in SEQ ID Nos. 6-15, or a nucleic acid complement thereof. In a related embodiment, the encoded amino acid sequence is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive amino acids up to the full length of one of SEQ ID NOS: 6-15. In yet another embodiment, the RING-membrane nucleic acid is an isolated nucleic acid encoding a polypeptide comprising a RING domain and a membrane domain.

In a further aspect, the invention provides a RING-receptor nucleic acid, such as an isolated nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence encoding a RING-receptor protein, such as a sequence of SEQ ID Nos: 57-72, or a sequence complementary thereto. In a related embodiment, the nucleic acid is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos.57-72, or a sequence complementary thereto or up to the full length of the gene of which said sequence is a fragment. In a further embodiment, the RING-receptor nucleic acid comprises a nucleic acid encoding an amino acid sequence as set forth in SEQ ID Nos. 16-31, or a nucleic acid complement thereof. In a

related embodiment, the encoded amino acid sequence is at least about 80%, 90%, 95%, or 97-98%, or 100% identical to a sequence corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive amino acids up to the full length of one of SEQ ID NOS: 16-31. In yet another embodiment, the RING-receptor nucleic acid is an isolated nucleic acid encoding a polypeptide comprising a RING domain and a receptor domain.

In one embodiment, the invention provides an expressible RING-SH3, RING-SH2, RING-membrane or RING-receptor nucleic acid operably linked to a transcriptional regulatory sequence, rendering the expressible nucleotide sequence suitable for use as an expression vector. In another embodiment, the nucleic acid may be included in an expression vector capable of replicating in a prokaryotic or eukaryotic cell. In a related embodiment, the invention provides a host cell transfected with the expression vector.

In yet another embodiment, the invention provides a substantially pure RING-SH3, RING-SH2, RING-membrane or RING-receptor nucleic acid which hybridizes under stringent conditions to a nucleic acid probe corresponding to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides up to the full length of one of SEQ ID Nos. 40-72, or a sequence complementary thereto or up to the full length of the gene of which said sequence is a fragment. The invention also provides an antisense oligonucleotide analog which hybridizes under stringent conditions to at least 12, at least 25, or at least 50 consecutive nucleotides of one of SEQ ID NOS 40-72, or a sequence complementary thereto.

In another embodiment, the invention provides a probe/primer comprising a substantially purified RING-SH3 oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos.40-44, or a sequence complementary thereto.

In another embodiment, the invention provides a probe/primer comprising a substantially purified RING-SH2 oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos.45-46, or a sequence complementary thereto.

In another embodiment, the invention provides a probe/primer comprising a substantially purified RING-membrane oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least about

12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos.47-56, or a sequence complementary thereto.

In another embodiment, the invention provides a probe/primer comprising a substantially purified RING-receptor oligonucleotide, said oligonucleotide containing a region of nucleotide sequence which hybridizes under stringent conditions to at least about 12, at least about 15, at least about 25, or at least about 40 consecutive nucleotides of sense or antisense sequence selected from SEQ ID Nos.57-72, or a sequence complementary thereto.

In preferred embodiments, a probe as described above selectively hybridizes with a target nucleic acid. In another embodiment, the probe may include a label group attached thereto and able to be detected. The label group may be selected from radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors. The invention further provides arrays of at least about 10, at least about 25, at least about 50, or at least about 100 different probes as described above attached to a solid support.

In another aspect, the invention provides polypeptides. In one embodiment, the invention pertains to a RING-SH3 polypeptide including an amino acid sequence encoded by a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 40-44, or a sequence complementary thereto, or a fragment comprising at least about 25, or at least about 40 amino acids thereof.

In another aspect, the invention provides polypeptides. In one embodiment, the invention pertains to a RING-SH2 polypeptide including an amino acid sequence encoded by a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 45-46, or a sequence complementary thereto, or a fragment comprising at least about 25, or at least about 40 amino acids thereof.

In another aspect, the invention provides polypeptides. In one embodiment, the invention pertains to a RING-membrane polypeptide including an amino acid sequence encoded by a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 47-56, or a sequence complementary thereto, or a fragment comprising at least about 25, or at least about 40 amino acids thereof.

In another aspect, the invention provides polypeptides. In one embodiment, the invention pertains to a RING-receptor polypeptide including an amino acid sequence encoded by a nucleic acid comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence of SEQ ID Nos. 57-72, or a sequence complementary thereto, or a fragment comprising at least about 25, or at least about 40 amino acids thereof.

In a preferred embodiment, the polypeptide is identical with or homologous to a RING-SH3, RING-SH2, RING-membrane or RING-receptor protein represented by SEQ ID Nos: 1-31. For instance, a polypeptide preferably has an amino acid sequence at least 70% homologous to a polypeptide represented by any of SEQ ID Nos: 1-31, though polypeptides with higher sequence homologies of, for example, 80%, 90% or 95% are also contemplated. The polypeptide can comprise a full length protein, such as represented in the sequence listings, or it can comprise a fragment of, for instance, at least 5, 10, 20, 50, 100, 150 or 200 amino acids in length.

In another preferred embodiment, the invention features a purified or recombinant polypeptide fragment of a RING-SH3, RING-SH2, RING-membrane or RING-receptor polypeptide, which polypeptide has the ability to modulate, e.g., mimic or antagonize, an activity of a wild-type RING-SH3, RING-SH2, RING-membrane or RING-receptor polypeptide. Preferably, the polypeptide fragment comprises a sequence identical or homologous to an amino acid sequence designated in one of SEQ ID Nos: 1-31.

Moreover, as described below, the RING-SH3, RING-SH2, RING-membrane or RING-receptor polypeptide can be either an agonist (e.g. mimics), or alternatively, an antagonist of a biological activity of a naturally occurring form of the protein, e.g., the polypeptide is able to modulate the intrinsic biological activity of a RING-SH3, RING-SH2, RING-membrane or RING-receptor complex, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, and the like.

The subject proteins can also be provided as chimeric molecules, such as in the form of fusion proteins. For instance, the AVMSF can be provided as a recombinant fusion protein which includes a second polypeptide portion, e.g., a second polypeptide having an amino acid sequence unrelated (heterologous) to the AVMSF, e.g. the second polypeptide portion is glutathione-S-transferase, e.g. the second polypeptide portion is an enzymatic activity such as alkaline phosphatase, e.g. the second polypeptide portion is an epitope tag, e.g. the second polypeptide is an affinity purification tag.

Yet another aspect of the present invention concerns an immunogen comprising an AVMSF in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for said AVMSF; e.g. a humoral response, e.g. an antibody response; e.g. a cellular response. In preferred embodiments, the immunogen comprising an antigenic determinant, e.g. a unique determinant, from a protein represented by one of SEQ ID Nos. 1-39.

In yet another aspect, this invention provides antibodies immunoreactive with one or more AVMSPs. In one embodiment, antibodies are specific for a RING domain, an SH3 domain, a SH2 domain, or a receptor domain and preferably the domain is part of an AVMSp. In a more specific embodiment, the domain is part of an amino acid sequence set forth in SEQ ID Nos. 1-39. In another embodiment, the antibodies are immunoreactive with one or more proteins having an amino acid sequence that is at least 80% identical to an amino acid sequence as set forth in SEQ ID Nos. 1-39. In other embodiments, an antibody is immunoreactive with one or more proteins having an amino acid sequence that is 85%, 90%, 95%, 98%, 99% or identical to an amino acid sequence as set forth in SEQ ID Nos. 1-39.

In an additional aspect, the invention provides complexes comprising an AVMSp and an AVMSp associated protein (an "AVMSp-AP"). In one embodiment, the invention provides an isolated protein complex comprising a RING-SH3, RING-SH2, RING-membrane or RING-receptor polypeptide in combination with at least one polypeptide selected from the group consisting of: a RING-SH3, a RING-SH2, a RING-membrane, a RING-receptor, a Gag protein, a Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin and a clathrin. In another embodiment, the isolated protein complex comprises a RING-SH3, RING-SH2, RING-membrane or RING-receptor polypeptide and a Gag protein in combination with a polypeptide selected from the group consisting of: a RING-SH3, a RING-SH2, a RING-membrane, a RING-receptor, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin and a clathrin.

In yet another embodiment, the invention provides an isolated protein complex comprising an AVMSp polypeptide and a HIV Gag protein in combination with a polypeptide selected from the group consisting of: RING-SH3, RING-SH2, RING-membrane, RING-receptor, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin and a clathrin. The invention also provides an isolated protein complex comprising a RING-SH3, RING-SH2, RING-membrane or RING-receptor polypeptide and an HIV Gag protein in combination with a polypeptide selected from the group consisting of: RING-SH3, RING-SH2, RING-membrane or RING-receptor, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin and a clathrin.

In yet another aspect, the invention provides an assay for screening test compounds for inhibitors, or alternatively, potentiators, of an interaction between an AVMSp and an AVMSp-AP. In the case of a RING-SH3 polypeptide, exemplary associated proteins

("RING-SH3-AP") include RING-SH3 proteins, E2 proteins (e.g. tsq101), Gag proteins, proteins comprising an L-domain, phosphatidylinositol-3-kinases, as well as proteins involved in endocytosis such as clathrins, actins, myosins, HSP60, HSP70, HSP90, STAM1, STAM2A, and STAM2B. An exemplary method includes the steps of (i) combining AVMS-AP (e.g. a RING-SH3-AP, RING-SH2-AP, RING-membrane-AP or RING-receptor-AP), an AVMS, and a test compound, e.g., under conditions (including, as desired, the addition of additional proteins) wherein, but for the test compound, the AVMS and an AVMS-AP are able to interact; and (ii) detecting the formation of a complex which includes the AVMS and an AVMS-AP. A statistically significant change, such as a decrease, in the formation of the complex in the presence of a test compound (relative to what is seen in the absence of the test compound) is indicative of a modulation, e.g., inhibition, of the interaction between the AVMS and an AVMS-AP. Similar assays may employ preformed AVMS-AVMS-AP complexes to assess the ability of the test compound to destabilize or stabilize the complex.

In yet another aspect, the invention provides cells carrying a recombinant form of an AVMS nucleic acid, often included on a vector. In further embodiments, cells carry a recombinant form of an AVMS nucleic acid and a recombinant form of a nucleic acid encoding a Gag protein and/or a polypeptide comprising an L domain motif, such as P(T/S)AP, PPxY or YxxL. In certain aspects, the cells are bacterial, and in other aspects the cells are eukaryotic cells, preferably a mammalian cell line.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art.

Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer

and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: Human PRT3 amino acid sequence (SEQ ID NO:1)

Figure 2A: N terminus protein fragment of hPRT3 (public gi:10432612; SEQ ID NO:2)

Figure 2B: C terminus protein fragment of hPRT3 (public gi:7959249; SEQ ID NO:39)

Figures 3 – 39 show amino acid sequences for examples of AVMSPP polypeptides.

Figures 40 – 78 show nucleic acid sequences encoding examples of AVMSPP polypeptides.

Figure 79: Human PRT3 full mRNA, annotated sequence

Figure 80: Domain analysis of human PRT3

Figure 81: Diagram of human PRT3 nucleic acids. The diagram shows the full-length PRT3 gene and the position of regions amplified by RT-PCR or targeted by siRNA used in figure 11.

Figure 82: Knockdown of PRT3 mRNA by siRNA duplexes. HeLa SS-6 cells were transfected with siRNA against Lamin A/C (lanes 1, 2) or PRT3 (lanes 3-10). PRT3 siRNA was directed against the coding region (153 - lanes 3,4; 155 - lanes 5,6) or the 3'UTR (157 - lanes 7, 8; 159 - lanes 9, 10). Cells were harvested 24 hours post-transfection, RNA extracted, and PRT3 mRNA levels compared by RT-PCR of a discrete sequence in the coding region of the PRT3 gene (see figure 10). GAPDH is used as an RT-PCR control in each reaction.

Figure 83: PRT3 affects the release of VLP from cells. A) Phosphorimages of SDS-PAGE gels of immunoprecipitations of 35S pulse-chase labeled Gag proteins are presented for cell and viral lysates from transfected HeLa cells that were either untreated or treated with PRT3 RNAi (50nM for 48 hours). The time during the chase period (1,2,3,4 and 5 after the pulse) are presented from left to right for each image.

Figure 84: Release of VLP from cells at steady state. HeLa cells were transfected with an HIV-encoding plasmid and siRNA. Lanes 1, 3 and 4 were transfected with wild-type HIV-encoding plasmid. Lane 2 was transfected with an HIV-encoding plasmid which contains a point mutation in p6 (PTAP to ATAP). Control siRNA (lamin A/C) was transfected to cells in lanes 1 and 2. siRNA to Tsg101 was transfected in lane 4 and siRNA to PRT3 in lane 3.

Figure 85: Exemplary HIV-1 Gag Nucleic Acid Sequence (Acc. No. NC_001802)

[SEQ ID NO:80]

Figure 86: Exemplary HIV-1 Gag Amino Acid Sequence (Acc. No. NP_057850)

[SEQ ID NO:81]

Figure 87: Exemplary HIV-1 p6 Amino Acid Sequence (SEQ ID NO:82)

Figure 88: Representative consensus terms for domains

Figure 89: Comparative Sequence Analysis – Amino Acid Grouping

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

The term “Alternate Viral Maturation Scaffolding Protein” or “AVMSP” is used herein to indicate a polypeptide comprising a first domain or functionality and a second domain. The first domain or functionality is selected from the group consisting of: an SH2 domain, an SH3 domain, a membrane spanning domain, and a receptor functionality (meaning the protein, or a portion thereof functions as a receptor). An AVMSP also comprises a second domain that is, for the purposes of this application, a RING domain. The first domain and second domain may be found in any order within the AVMSP sequence (i.e. the first domain need not be N-terminal to the second domain). The term AVMSP encompasses proteins represented as SEQ ID Nos. 1-39. It is understood that polypeptides having functional roles, such as receptors, may further comprise membrane or other domains. In other embodiments, AVMSPs may further comprise one or more C2 domains. In cells infected with viruses that utilize a Gag-dependent pathway for budding and release, an AVMSP may act to assemble complexes of proteins that mediate release. AVMSP complexes may stimulate ubiquitylation of certain proteins or stimulate membrane fusion or both. Any single AVMSP may form multiple different AVMSP complexes at different times.

The term “Alternate Viral Maturation Scaffolding Protein-Associated Protein” (AVMSP-AP) refers to protein capable of interacting and/or binding to an AVMSP. Generally, the AVMSP-AP may interact directly or indirectly with the AVMSP. Examples of these proteins include for example the “Late domain” or “L domain”, which is a small portion of a Gag protein that promotes efficient release of virion particles from the membrane of the host cell. L domains typically comprise one or more short motifs (L motifs). Exemplary sequences include: P(T/S)AP, PxxL, PPxY (eg. PPPY), YxxL (eg. YPDL), PxxP. Other exemplary AVMSP-APs are provided throughout.

The term "binding" refers to a direct association between two molecules, due to, for example, covalent, electrostatic, hydrophobic, ionic and/or hydrogen-bond interactions under physiological conditions.

A "C2 domain" is a calcium binding domain. Certain C2 domains comprise the consensus sequence set forth in Figure 14 as "Consensus/80%". Other C2 domains comprise the consensus sequences set forth as "Consensus/65%" or "Consensus/50%".

"Cells," "host cells" or "recombinant host cells" are terms used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding a polypeptide with a second amino acid sequence defining a domain foreign to and not substantially homologous with any domain of the first amino acid sequence. A chimeric protein may present a foreign domain which is found (albeit in a different protein) in an organism which also expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms.

The terms "compound", "test compound" and "molecule" are used herein interchangeably and are meant to include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, natural product extract libraries, and any other molecules (including, but not limited to, chemicals, metals and organometallic compounds).

The phrase "conservative amino acid substitution" refers to grouping of amino acids on the basis of certain common properties. A functional way to define common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). According to such analyses, groups of amino acids may be defined where amino acids within a group exchange preferentially with each other, and therefore resemble each other most in their impact on the overall protein structure (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). Examples of amino acid groups defined in this manner include:

- (i) a charged group, consisting of Glu and Asp, Lys, Arg and His,
- (ii) a positively-charged group, consisting of Lys, Arg and His,
- (iii) a negatively-charged group, consisting of Glu and Asp,

- (iv) an aromatic group, consisting of Phe, Tyr and Trp,
- (v) a nitrogen ring group, consisting of His and Trp,
- (vi) a large aliphatic nonpolar group, consisting of Val, Leu and Ile,
- (vii) a slightly-polar group, consisting of Met and Cys,
- (viii) a small-residue group, consisting of Ser, Thr, Asp, Asn, Gly, Ala, Glu, Gln and Pro,
- (ix) an aliphatic group consisting of Val, Leu, Ile, Met and Cys, and
- (x) a small hydroxyl group consisting of Ser and Thr.

In addition to the groups presented above, each amino acid residue may form its own group, and the group formed by an individual amino acid may be referred to simply by the one and/or three letter abbreviation for that amino acid commonly used in the art.

A "conserved residue" is an amino acid that is relatively invariant across a range of similar proteins. Often conserved residues will vary only by being replaced with a similar amino acid, as described above for "conservative amino acid substitution".

The term "domain" as used herein refers to a region of a protein that comprises a particular structure and/or performs a particular function.

The term "Gag protein" or "Gag polypeptide" refers to a polypeptide having Gag activity and preferably comprising an L (or late) domain. Exemplary Gag proteins include a motif such as PXXP, PPXY, PXXY, YXXL, RXXPPXP, RPDPTAP, RPLPVAP, RPEPTAP, PTAPPEY, PTAPPEE and/or RPEPTAPPEE. An exemplary HIV-1 Gag protein is SEQ ID NO: 32. Typically, an HIV Gag protein comprises a p6 protein.

"Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology and identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site occupied by the same or a similar amino acid residue (e.g., similar in steric and/or electronic nature), then the molecules can be referred to as homologous (similar) at that position. Expression as a percentage of homology/similarity or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. A sequence which is "unrelated" or "non-homologous" shares less than 40% identity, though preferably less than 25% identity with a sequence of the present invention. In comparing two sequences, the absence of residues (amino acids or nucleic acids) or presence of extra residues also decreases the identity and homology/similarity.

The term "homology" describes a mathematically based comparison of sequence similarities which is used to identify genes or proteins with similar functions or motifs. The nucleic acid and protein sequences of the present invention may be used as a "query sequence" to perform a search against public databases to, for example, identify other family members, related sequences or homologs. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) *J Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and BLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

As used herein, "identity" means the percentage of identical nucleotide or amino acid residues at corresponding positions in two or more sequences when the sequences are aligned to maximize sequence matching, i.e., taking into account gaps and insertions. Identity can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S. F. et al., *J. Molec. Biol.* 215: 403-410 (1990) and Altschul et al. *Nuc. Acids Res.* 25: 3389-3402 (1997)). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md.

20894; Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

The term "including" is used herein to mean, and is used interchangeably with, the phrase "including but not limited to".

The term "intron" refers to a portion of nucleic acid that is initially transcribed into RNA but later removed such that it is not, for the most part, represented in the processed mRNA. Intron removal occurs through reactions at the 5' and 3' ends, typically referred to as 5' and 3' splice sites, respectively. Alternate use of different splice sites results in splice variants. An intron is not necessarily situated between two "exons", or portions that code for amino acids, but may instead be positioned, for example, between the promoter and the first exon. An intron may be self-splicing or may require cellular components to be spliced out of the mRNA. A "heterologous intron" is an intron that is inserted into a coding sequence that is not naturally associated with that coding sequence. In addition, a heterologous intron may be a generally natural intron wherein one or both of the splice sites have been altered to provide a desired quality, such as increased or decreased splice efficiency. Heterologous introns are often inserted, for example, to improve expression of a gene in a heterologous host, or to increase the production of one splice variant relative to another. As an example, the rabbit beta-globin gene may be used, and is commercially available on the pCI vector from Promega Inc. Other exemplary introns are provided in Lacy-Hulbert et al. (2001) Gene Ther 8(8):649-53.

The term "isolated", as used herein with reference to the subject proteins and protein complexes, refers to a preparation of protein or protein complex that is essentially free from contaminating proteins that normally would be present with the protein or complex, e.g., in the cellular milieu in which the protein or complex is found endogenously. Thus, an isolated protein complex is isolated from cellular components that normally would "contaminate" or interfere with the study of the complex in isolation, for instance while screening for modulators thereof. It is to be understood, however, that such an "isolated" complex may incorporate other proteins the modulation of which, by the subject protein or protein complex, is being investigated.

The term "isolated" as also used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules in a form which does not occur in nature. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

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-stranded (such as sense or antisense) and double-stranded polynucleotides.

The term "maturation" as used herein refers to the processing of viral proteins leading to the pinching off of nascent virion from the cell membrane, including, for example, assembly, budding and release.

A "membrane associated protein" is meant to include proteins that are integral membrane proteins as well as proteins that are stably associated with a membrane.

A "membrane spanning domain" or "transmembrane domain" or membrane domain" is a domain that traverses a biological lipid bilayer membrane from one side to the other. Generally membrane domains are identified as a region of a protein having a high hydrophobicity. Membrane domains are typically between 15 and 25 amino acids in length. Exemplary methods for identifying membrane domains are provided in Figure 129.

The term "or" is used herein to mean, and is used interchangeably with, the term "and/or", unless context clearly indicates otherwise.

The term "p6" or p6gag" is used herein to refer to an HIV protein comprising a viral L domain. Antibodies that bind to a p6 domain are referred to as "anti-p6 antibodies". p6 also refers to proteins that comprise artificially engineered L domains including, for example, L domains comprising a series of L motifs. An exemplary HIV-1 p6 is SEQ ID NO: 85.

A "profile" is used herein to indicate an aggregate of information regarding a preparation of cell or membrane surface proteins. A profile will comprise, at minimum, information regarding the presence or absence of such proteins. More typically, a profile will comprise information regarding the presence or absence of a plurality of such proteins. In addition, a profile may contain other information about each identified protein, such as relative or absolute amount of protein present, the degree of post-translational modification, membrane topology, three-dimensional structure, isoelectric point, molecular weight, etc. A "test profile" is a profile obtained from a subject of unknown diagnostic state. A "reference profile" is a profile obtained from subject known to be infected or uninfected.

The terms peptides, proteins and polypeptides are used interchangeably herein.

The term "purified protein" refers to a preparation of a protein or proteins which are preferably isolated from, or otherwise substantially free of, other proteins normally associated with the protein(s) in a cell or cell lysate. The term "substantially free of other cellular

proteins" (also referred to herein as "substantially free of other contaminating proteins") is defined as encompassing individual preparations of each of the component proteins comprising less than 20% (by dry weight) contaminating protein, and preferably comprises less than 5% contaminating protein. Functional forms of each of the component proteins can be prepared as purified preparations by using a cloned gene as described in the attached examples. By "purified", it is meant, when referring to component protein preparations used to generate a reconstituted protein mixture, that the indicated molecule is present in the substantial absence of other biological macromolecules, such as other proteins (particularly other proteins which may substantially mask, diminish, confuse or alter the characteristics of the component proteins either as purified preparations or in their function in the subject reconstituted mixture). The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 85% by weight, more preferably 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above.

A "receptor" or "protein having a receptor function" is a protein that interacts with an extracellular ligand or a ligand that is within the cell but in a space that is topologically equivalent to the extracellular space (eg. inside the Golgi, inside the endoplasmic reticulum, inside the nuclear membrane, inside a lysosome or transport vesicle, etc.). Exemplary receptors are identified herein by annotation as such in various public databases. Receptors often have membrane domains.

A "recombinant nucleic acid" is any nucleic acid that has been placed adjacent to another nucleic acid by recombinant DNA techniques. A "recombined nucleic acid" also includes any nucleic acid that has been placed next to a second nucleic acid by a laboratory genetic technique such as, for example, transformation and integration, transposon hopping or viral insertion. In general, a recombined nucleic acid is not naturally located adjacent to the second nucleic acid.

The term "recombinant protein" refers to a protein of the present invention which is produced by recombinant DNA techniques, wherein generally DNA encoding the expressed protein is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase "derived from", with respect to a recombinant gene encoding the recombinant protein is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native

protein, or an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions of a naturally occurring protein.

A "RING domain" or "Ring Finger" is a zinc-binding domain with a defined octet of cysteine and histidine residues. Certain RING domains comprise the consensus sequences as set forth below (amino acid nomenclature is as set forth in Table 1): Cys Xaa Xaa Cys Xaa₁₀₋₂₀ Cys Xaa His Xaa₂₋₅ Cys Xaa Xaa Cys Xaa₁₃₋₅₀ Cys Xaa Xaa Cys or Cys Xaa Xaa Cys Xaa₁₀₋₂₀ Cys Xaa His Xaa₂₋₅ His Xaa Xaa Cys Xaa₁₃₋₅₀ Cys Xaa Xaa Cys. Preferred RING domains of the invention bind to various protein partners to form a complex that has ubiquitin ligase activity. RING domains preferably interact with at least one of the following protein types: F box proteins, E2 ubiquitin conjugating enzymes and cullins.

A "scaffolding protein" is a protein that brings together two or more different proteins that interact to accomplish one or more particular functions. A scaffolding protein may, in addition to acting as a scaffold, carry out biochemical functions on its own or as part of a complex.

"Small molecule" as used herein, is meant to refer to a composition, which has a molecular weight of less than about 5 kD and most preferably less than about 2.5 kD. Small molecules can be nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic (carbon containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures comprising arrays of small molecules, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the invention.

An "SH2" or "Src Homology 2" domain is a protein domain of generally about 100 amino acid residues. SH2 domains function as regulatory modules of intracellular signalling cascades by interacting with high affinity to phosphotyrosine-containing target peptides in a sequence-specific and phosphorylation-dependent manner.

An "SH3" or "Src Homology 3" domain is a protein domain of generally about 60 amino acid residues first identified as a conserved sequence in the non-catalytic part of several cytoplasmic protein tyrosine kinases (e.g. Src, Abl, Lck). SH3 domains mediate assembly of specific protein complexes via binding to proline-rich peptides.

As used herein, the term "specifically hybridizes" refers to the ability of a nucleic acid probe/primer of the invention to hybridize to at least 12, 15, 20, 25, 30, 35, 40, 45, 50 or 100 consecutive nucleotides of a target gene sequence, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it has less than 15%, preferably less than 10%, and more preferably less than 5% background hybridization to a cellular nucleic acid (e.g.,

mRNA or genomic DNA) other than the target gene. A variety of hybridization conditions may be used to detect specific hybridization, and the stringency is determined primarily by the wash stage of the hybridization assay. Generally high temperatures and low salt concentrations give high stringency, while low temperatures and high salt concentrations give low stringency. Low stringency hybridization is achieved by washing in, for example, about 2.0 x SSC at 50 °C, and high stringency is achieved with about 0.2 x SSC at 50 °C. Further descriptions of stringency are provided below.

“STAM” proteins include a family of proteins involved in receptor mediated exo- and endocytosis as well as cellular signalling, generally. STAM proteins generally comprise an N-terminal VHS homology domain, a ubiquitin-interacting motif and an SH3 domain and optionally an immunoreceptor tyrosine-based activation motif. STAM 1 and STAM 2A are involved in cytokine-mediated signalling for DNA synthesis and c-myc induction. EAST and STAM 2A/Hbp play a role in receptor-mediated endo- and exocytosis and probably also in the regulation of actin cytoskeleton. (Lohi et al. FEBS Lett 2001 Nov 23;508(3):287-90)

As applied to polypeptides, “substantial sequence identity” means that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap which share at least 90 percent sequence identity, preferably at least 95 percent sequence identity, more preferably at least 99 percent sequence identity or more. Preferably, residue positions which are not identical differ by conservative amino acid substitutions. For example, the substitution of amino acids having similar chemical properties such as charge or polarity are not likely to effect the properties of a protein. Examples include glutamine for asparagine or glutamic acid for aspartic acid.

“Transcriptional regulatory sequence” is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of a recombinant protein gene is under the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring form of the protein.

A “UIM” domain is a ubiquitin binding motif.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. 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". 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.

A "virion" is a complete viral particle; nucleic acid and capsid (and a lipid envelope in some viruses).

A "VHS" domain is a "Vps27p, Hrs and STAM" domain, named for the proteins in which it has been identified, and includes a DXXLL sequence motif. VHS domains have also been identified in the GGA (Golgi-localized, gamma-ear-containing, ADP-ribosylation-factor-binding) proteins. In certain embodiments, VHS domains of the invention recognize one or more acidic-cluster-dileucine signals found in the cytoplasmic tails of sorting receptors, such as mannose-6-phosphate receptors. (Misra et al. (2002) Nature 2002 Feb 21;415(6874):933-7)

Table 1: Abbreviations for classes of amino acids*

Symbol	Category	Amino Acids Represented
X1	Alcohol	Ser, Thr
X2	Aliphatic	Ile, Leu, Val
Xaa	Any	Ala, Cys, Asp, Glu, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, Tyr

X4	Aromatic	Phe, His, Trp, Tyr
X5	Charged	Asp, Glu, His, Lys, Arg
X6	Hydrophobic	Ala, Cys, Phe, Gly, His, Ile, Lys, Leu, Met, Thr, Val, Trp, Tyr
X7	Negative	Asp, Glu
X8	Polar	Cys, Asp, Glu, His, Lys, Asn, Gln, Arg, Ser, Thr
X9	Positive	His, Lys, Arg
X10	Small	Ala, Cys, Asp, Gly, Asn, Pro, Ser, Thr, Val
X11	Tiny	Ala, Gly, Ser
X12	Turnlike	Ala, Cys, Asp, Glu, Gly, His, Lys, Asn, Gln, Arg, Ser, Thr
X13	Asparagine-Aspartate	Asn, Asp

* Abbreviations as adopted from http://smart.embl-heidelberg.de/SMART_DATA/alignments/consensus/grouping.html.

2. Overview

In certain aspects, the invention relates to the observation that AVMSM polypeptides are involved in viral maturation process such as assembly, budding and/or release. Any one AVMSM may be involved at one or more stages of viral maturation and may form one or more complexes with viral and/or host proteins.

In certain aspects, AVMSM sequences disclosed herein, and related methods and compositions may be used to manipulate a variety of significant biological processes. In certain embodiments, AVMSM nucleic acids and polypeptides of the invention are involved in a viral lifecycle. In one embodiment, an SH3 domain of a RING-SH3 protein binds to a viral Gag sequence represented by a consensus sequence P(T/S)AP, RXXP(T/S)AP or

PXXDY, such as, for example, the PTAP or PFRDY sequences of HIV Gag (positioned, for example, at 455-458 and 292-296, respectively).

In a further embodiment, AVMSP polypeptides may be involved in the formation of endocytosis-like or vesicle trafficking-like complexes that are involved in a stage of viral maturation. Such complexes may include complexes similar to those formed in connection with clathrin-mediated vesicle trafficking or those formed in connection with coatamer-coated vesicle trafficking.

Viral assembly, budding and release is expected to require a range of different protein complexes that incorporate host proteins involved in different aspects of vesicle trafficking, including vesicle formation proteins such as ARFs, COPs, RABs and clathrins, cytoskeletal proteins such as actins and myosins, cytoskeletal regulators such as Rac and Rho, heat shock proteins, STAM proteins and viral proteins, particularly proteins having a Late domain, such as Gag. Such complexes are expected to incorporate one or more AVMSPs, and particularly AVMSPs such as RING-SH3 proteins. At each stage of vesicle transport, the exact components of the relevant complexes may shift, but in general, it is understood that disrupting the ability of an AVMSP to participate in complex formation or dissolution will be effective in disrupting viral production. In certain embodiments, it is understood that interfering with AVMSP activity will not decrease the rate of viral production but will result in the production of defective viral particles having decreased ability to infect another host cell.

It is generally understood that AVMSPs, especially AVMSPs having similar functional domains, may exhibit substantial functional overlap in both native host functions and in the viral lifecycle. Accordingly, the invention provides methods for inhibiting a virus by interfering with the activity of more than one AVMSP. A preferred antiviral agent is able, as a single compound or mixture of compounds to interfere with more than one AVMSP.

3. Exemplary Nucleic Acids and Expression Vectors

In certain aspects the invention provides nucleic acids encoding Alternate Viral Maturation Scaffolding Proteins (AVMSPs), such as for example RING-SH3 proteins, RING-SH2 proteins, RING-membrane proteins and RING-receptor proteins. There are four basic classes of AVMSPs: the RING-SH3 class, comprising at least one SH3 domain and at least one RING domain, the RING-SH2 class, comprising at least one SH2 domain and at least one RING domain, the RING-membrane class, comprising at least one RING domain and at least one transmembrane domain, and the RING-receptor class, comprising at least one

RING domain and having receptor functionality. These four classes are not mutually exclusive, as, for example, a polypeptide may comprise a RING domain, and SH3 domain and an SH2 domain. In preferred embodiments, proteins of any of the four classes comprise at least one C2 domain.

Nucleic acids of the invention is further understood to include nucleic acids that encode variants of AVMSPs. Variant nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include coding sequences that differ from the nucleotide sequence of the coding sequence designated in Tables 2-4 e.g., due to the degeneracy of the genetic code. Variants will 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 nucleotide sequence of a coding sequence designated in Tables 2 and 3. Alternatively put, variants will also include nucleotide sequences that hybridize under moderately stringent conditions, for example at about 2.0 x SSC and about 40 °C to the nucleotide sequence of a coding sequence designated in Tables 2-4. In another embodiment, equivalent nucleic acid sequences include sequences that will hybridize under highly stringent conditions to a nucleotide sequence of a coding sequence designated in Tables 2-4.

One of ordinary skill in the art will understand readily that appropriate stringency conditions which promote DNA hybridization can be varied. For example, one could perform the hybridization at 6.0 x sodium chloride/sodium citrate (SSC) at about 45 °C, followed by a wash of 2.0 x SSC at 50 °C. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50 °C to a high stringency of about 0.2 x SSC at 50 °C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22 °C, to high stringency conditions at about 65 °C. Both temperature and salt may be varied, or temperature or salt concentration may be held constant while the other variable is changed. In one embodiment, the invention provides nucleic acids which hybridize under low stringency conditions of 6 x SSC at room temperature followed by a wash at 2 x SSC at room temperature.

In one embodiment, variants will further include nucleic acid sequences derived from and evolutionarily related to a nucleotide sequence of a coding sequence designated in Tables 2-4.

Table 2: Exemplary AVMSPs

<u>Name</u>	<u>Accession Nos.</u> <u>Nucleotide; Protein</u>	<u>Nucleotide</u> <u>SEQ I.D. No.</u>	<u>Protein</u> <u>SEQ I.D. No.</u>	<u>Selected Domains or</u> <u>Features of Interest</u>
Human PRT3 Coding sequence		40	1	RING, SH3
Human PRT3 cDNA		41	1	
5' cDNA fragment of human PRT3		42	2	
3' mRNA fragment of human PRT3		43	39	
SIMILAR TO MUS MUSCULUS PLENTY OF SH3S (POSH)	EMBL; AK021429; BAB13822.1	44	3	RING, SH3
SH3-BINDING PROTEIN CBL-B	EMBL; U26710; AAB09291.1	45	4	RING, SH2
SH3-BINDING PROTEIN CBL-C	EMBL; AB028645; BAA86298.1	46	5	RING, SH2
MULTIPLE MEMBRANE SPANNING RECEPTOR TRC8	EMBL; AF064801; AAC39930.1	47	6	RING, Membrane
AUTOCRINE MOTILITY FACTOR RECEPTOR	EMBL; AF124145; AAD56722.1	48	7	RING, Membrane
MEMBRANE- ASSOCIATED NUCLEIC ACID BINDING PROTEIN (FRAGMENT)	EMBL; AF255303; AAG00432.1	49	8	RING, Membrane
NADH- UBIQUINONE OXIDOREDUCTASE MWFE SUBUNIT	EMBL; X81900; CAA57489.1	50	9	RING, Membrane
PEROXISOME ASSEMBLY FACTOR-1 (PAF-1)	EMBL; M86852; AAC12785.1	51	10	RING, Membrane
PEROXISOME ASSEMBLY PROTEIN 10	EMBL; AF060502; AAC18133.1	52	11	RING, Membrane
PEROXISOME ASSEMBLY PROTEIN 12	EMBL; U91521; AAC68812.1	53	12	RING, Membrane
ACETYLCHOLINE RECEPTOR- ASSOCIATED 43	EMBL; Z33905; CAA83954.1	54	13	RING, Membrane

KDA PROTEIN				
PEROXISOMAL MEMBRANE PROTEIN 3	EMBL; BC000661; AAH00661.1	55	14	RING, Membrane
HsRmal	EMBL; AB056869; BAB39359.1	56	15	RING, Membrane
HTRIP	EMBL; U77845; AAB52993.1	57	16	RING, Receptor
CYSTEIN RICH DOMAIN ASSOCIATED TO RING AND TRAF PROTEIN	EMBL; X80200; CAA56491.1	58	17	RING, Receptor
PML-RAR	EMBL; M73779; AAA60126.1	59	18	RING, Receptor
TRAF5	EMBL; AB000509; BAA25262.1	60	19	RING, Receptor
TUMOR NECROSIS FACTOR RECEPTOR- ASSOCIATED FACTOR 4A	EMBL; AF082185; AAC32376.1	61	20	RING, Receptor
HFB30	EMBL; AB022663; BAA78677.1	62	21	RING, Receptor
BIR2	EMBL; L49432; AAC41943.1	63	22	RING, Receptor
BIR3	EMBL; L49431; AAC41942.1	64	23	RING, Receptor
CBL	EMBL; X57110; CAA40393.1	65	24	RING, Receptor
PML	EMBL; M79462; AAA60388.1	66	25	RING, Receptor
RAG1	EMBL; M29474; AAA60248.1	67	26	RING, Receptor
RAPSYN	EMBL; Z33905; CAA83954.1	68	27	RING, Receptor
TIFI-ALPHA	EMBL; AF009353; AAB63585.1	69	28	RING, Receptor
TIFI-GAMMA	EMBL; AF119043; AAD17259.1	70	29	RING, Receptor
TRAF2	EMBL; U12597;	71	30	RING, Receptor

	AAA87706.1			
TRAF3	EMBL; U21092; AAC50112.1	72	31	RING, Receptor
HIP116	EMBL; L34673; AAA67436.1	73	32	RING
STAF50	EMBL; X82200; CAA57684.1	74	33	RING
HLTF-1	EMBL; Z46606; CAA86571.1	75	34	RING
BIR4	EMBL; U45880; AAC50373.1	76	35	RING
ADENOVIRUS 5 E1A-BINDING PROTEIN	EMBL; X86098; CAA60052.1	77	36	RING
HT2A	EMBL; U18543; AAA86474.1	78	37	RING
TRAF3	EMBL; U21092; AAC50112.1	79	38	RING

Table 3: Exemplary POSH nucleic acids

<u>Sequence Name</u>	<u>Organism</u>	<u>Accession Number</u>
cDNA FLJ11367 fis, clone HEMBA1000303	Homo sapiens	AK021429
Plenty of SH3 domains (POSH) mRNA	Mus musculus	NM_021506
Plenty of SH3s (POSH) mRNA	Mus musculus	AF030131
Plenty of SH3s (POSH) mRNA	Drosophila melanogaster	NM_079052
Plenty of SH3s (POSH) mRNA	Drosophila melanogaster	AF220364

Table 4: Exemplary POSH polypeptides

<u>Sequence Name</u>	<u>Organism</u>	<u>Accession Number</u>
SH3 domains- containing protein POSH	Mus musculus	T09071
plenty of SH3 domains	Mus musculus	NP_067481
Plenty of SH3s; POSH	Mus musculus	AAC40070
Plenty of SH3s	Drosophila melanogaster	AAF37265
LD45365p	Drosophila melanogaster	AAK93408
POSH gene product	Drosophila melanogaster	AAF57833
Plenty of SH3s	Drosophila melanogaster	NP_523776

Isolated nucleic acids which differ from the nucleotide sequences encoding a protein designated in Tables 2 - 4 due to degeneracy in the genetic code are also within the scope of the invention. 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 are synonyms for histidine) may result in "silent" mutations which do not affect the amino acid sequence of the protein. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject proteins will exist among mammalian cells. 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 a particular protein may exist among individuals of a given species due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of this invention.

Another aspect of the invention relates to the use of the isolated nucleic acid in "antisense" therapy. As used herein, antisense therapy refers to administration or *in situ* generation of oligonucleotide probes or their derivatives which specifically hybridize (e.g. binds) under cellular conditions with the cellular mRNA and/or genomic DNA encoding one of the subject AVMSPs so as to inhibit expression of that protein, e.g. by inhibiting transcription and/or translation. The binding 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 which relies on specific binding to oligonucleotide sequences.

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 an AVMSP. Alternatively, the antisense construct is an oligonucleotide probe which 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 an AVMSP. Such oligonucleotide probes are preferably modified oligonucleotide which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and is 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. Patents 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

Accordingly, the modified oligomers of the invention are useful in therapeutic, diagnostic, and research contexts. In therapeutic applications, the oligomers are utilized in a manner appropriate for antisense therapy in general.

In addition to use in therapy, the oligomers of the invention may be used as diagnostic reagents to detect the presence or absence of the target DNA or RNA sequences to which they specifically bind, such as for determining the level of expression of a gene of the invention or for determining whether a gene of the invention contains a genetic lesion.

In another aspect of the invention, the subject nucleic acid is provided in an expression vector comprising a nucleotide sequence encoding a subject AVMSP polypeptide and operably linked to at least one regulatory sequence. Operably linked is intended to mean that the nucleotide sequence is linked to a regulatory sequence in a manner which allows expression of the nucleotide sequence. Regulatory sequences are art-recognized and are selected to direct expression of the polypeptide having an activity of an AVMSP. Accordingly, the term regulatory sequence includes promoters, enhancers and other expression control elements. Exemplary regulatory sequences are described in Goeddel; *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, CA (1990). For instance, any of a wide variety of expression control sequences that control the expression of a DNA sequence when operatively linked to it may be used in these vectors to express DNA sequences encoding an AVMSP. Such useful expression control sequences,

include, for example, the early and late promoters of SV40, tet promoter, adenovirus or cytomegalovirus immediate early promoter, the lac system, the trp system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage lambda, the control regions for fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. Moreover, the vector's copy number, the ability to control that copy number and the expression of any other protein encoded by the vector, such as antibiotic markers, should also be considered.

As will be apparent, the subject gene constructs can be used to cause expression of the subject AVMSF polypeptides in cells propagated in culture, e.g. to produce proteins or polypeptides, including fusion proteins or polypeptides, for purification.

This invention also pertains to a host cell transfected with a recombinant gene including a coding sequence for one or more of the subject AVMSF. The host cell may be any prokaryotic or eukaryotic cell. For example, a polypeptide of the present invention may be expressed in bacterial cells such as *E. coli*, insect cells (e.g., using a baculovirus expression system), yeast, or mammalian cells. Other suitable host cells are known to those skilled in the art.

Accordingly, the present invention further pertains to methods of producing the subject AVMSF polypeptides. For example, a host cell transfected with an expression vector encoding an AVMSF polypeptide can be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the polypeptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins, including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for particular epitopes of the

polypeptide. In a preferred embodiment, the AVMSP is a fusion protein containing a domain which facilitates its purification, such as an AVMSP-GST fusion protein, AVMSP-cellulose binding domain fusion protein, etc.

A nucleotide sequence encoding an AVMSP can be used to produce a recombinant form of the protein via microbial or eukaryotic cellular processes. Ligating the polynucleotide sequence into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial) cells, are standard procedures.

A recombinant AVMSP can be produced by ligating the cloned gene, or a portion thereof, into a vector suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vehicles for production of a recombinant AVMSP include plasmids and other vectors. For instance, suitable vectors for the expression of an AVMSP include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al., (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein). These vectors can replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used.

The preferred mammalian expression vectors contain both prokaryotic sequences to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papilloma virus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. Examples of other viral (including retroviral) expression systems can be found below in the description of gene therapy delivery systems.

The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press, 1989) Chapters 16 and 17. In some instances, it may be desirable to express the recombinant AVMSF by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the β -gal containing pBlueBac III).

It is well known in the art that a methionine at the N-terminal position can be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli* (Ben-Bassat et al., (1987) *J. Bacteriol.* 169:751-757) and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins (Miller et al., (1987) *PNAS USA* 84:2718-1722). Therefore, removal of an N-terminal methionine, if desired, can be achieved either *in vivo* by expressing such recombinant polypeptides in a host which produces MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP (e.g., procedure of Miller et al.).

Alternatively, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable, e.g., to produce an immunogenic fragment of an AVMSF. For example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of the AVMSF to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising a portion of the protein as part of the virion. The Hepatitis B surface antigen can also be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of an AVMSF and the poliovirus capsid protein can be created to enhance immunogenicity (see, for example, EP Publication NO: 0259149; and Evans et al., (1989) *Nature* 339:385; Huang et al., (1988) *J. Virol.* 62:3855; and Schlienger et al., (1992) *J. Virol.* 66:2).

The Multiple Antigen Peptide system for peptide-based immunization can be utilized, wherein a desired portion of an AVMSF is obtained directly from organo-chemical synthesis

of the peptide onto an oligomeric branching lysine core (see, for example, Posnett et al., (1988) *JBC* 263:1719 and Nardelli et al., (1992) *J. Immunol.* 148:914). Antigenic determinants of an AVMSp can also be expressed and presented by bacterial cells.

In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the expressed fusion protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified AVMSp (e.g., see Hochuli et al., (1987) *J. Chromatography* 411:177; and Janknecht et al., *PNAS USA* 88:8972).

Techniques for making fusion genes are well known. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., John Wiley & Sons: 1992).

4. Exemplary Polypeptides

The present invention also makes available isolated and/or purified forms of the subject AVMSps, which are isolated from, or otherwise substantially free of, other intracellular proteins which might normally be associated with the protein or a particular complex including the protein. In certain embodiments, polypeptides of the invention have an amino acid sequence that is at least 60% identical to an amino acid sequence as set forth in SEQ ID Nos. 1-39. In other embodiments, the polypeptide has an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% or 100% identical to an amino acid sequence as set forth in SEQ ID Nos. 1-39.

In another aspect, the invention provides polypeptides that are agonists or antagonists of AVMSps. Variants and fragments of an AVMSp may have a hyperactive or constitutive

activity, or, alternatively, act to prevent AVMSPs from performing one or more functions. For example, a truncated form lacking one or more domain may have a dominant negative effect.

Another aspect of the invention relates to polypeptides derived from a full-length AVMSp. Isolated peptidyl portions of the subject proteins can be obtained by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such polypeptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, any one of the subject proteins can be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of the formation of a specific protein complex, or more generally of an AVMSp, such as by microinjection assays.

It is also possible to modify the structure of the subject AVMSps for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., ex vivo shelf life and resistance to proteolytic degradation in vivo). Such modified polypeptides, when designed to retain at least one activity of the naturally-occurring form of the protein, are considered functional equivalents of the AVMSps described in more detail herein. Such modified polypeptides can be produced, for instance, by amino acid substitution, deletion, or addition.

For instance, it is reasonable to expect, for example, that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e. conservative mutations) will not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan;

(5) amide = asparagine, glutamine; and (6) sulfur -containing = cysteine and methionine. (see, for example, Biochemistry, 2nd ed., Ed. by L. Stryer, W.H. Freeman and Co., 1981). Whether a change in the amino acid sequence of a polypeptide results in a functional homolog can be readily determined by assessing the ability of the variant polypeptide to produce a response in cells in a fashion similar to the wild-type protein. For instance, such variant forms of an AVMSF can be assessed, e.g., for their ability to bind to another polypeptide, e.g., another AVMSF or another protein involved in viral maturation. Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This invention further contemplates a method of generating sets of combinatorial mutants of the subject AVMSFs, as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs) that are functional in binding to an AVMSF. The purpose of screening such combinatorial libraries is to generate, for example, AVMSF homologs which can act as either agonists or antagonist, or alternatively, which possess novel activities all together. Combinatorially-derived homologs can be generated which have a selective potency relative to a naturally occurring AVMSF. Such proteins, when expressed from recombinant DNA constructs, can be used in gene therapy protocols.

Likewise, mutagenesis can give rise to homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered protein can be rendered either more stable or less stable to proteolytic degradation or other cellular process which result in destruction of, or otherwise inactivation of the AVMSF of interest. Such homologs, and the genes which encode them, can be utilized to alter AVMSF expression by modulating the half-life of the protein. For instance, a short half-life can give rise to more transient biological effects and, when part of an inducible expression system, can allow tighter control of recombinant AVMSF levels within the cell. As above, such proteins, and particularly their recombinant nucleic acid constructs, can be used in gene therapy protocols.

In similar fashion, AVMSF homologs can be generated by the present combinatorial approach to act as antagonists, in that they are able to interfere with the ability of the corresponding wild-type protein to function.

In a representative embodiment of this method, the amino acid sequences for a population of AVMSF homologs are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, homologs from one or more species, or homologs from the same species but which differ due to mutation. Amino acids

which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. In a preferred embodiment, the combinatorial library is produced by way of a degenerate library of genes encoding a library of polypeptides which each include at least a portion of potential AVMSP sequences. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential AVMSP nucleotide sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display).

There are many ways by which the library of potential homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then be ligated into an appropriate gene for expression. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential AVMSP sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al., (1981) *Recombinant DNA, Proc. 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al., (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al., (1984) *Science* 198:1056; Ike et al., (1983) *Nucleic Acid Res.* 11:477). Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al., (1990) *Science* 249:386-390; Roberts et al., (1992) *PNAS USA* 89:2429-2433; Devlin et al., (1990) *Science* 249: 404-406; Cwirla et al., (1990) *PNAS USA* 87: 6378-6382; as well as U.S. Patent Nos: 5,223,409, 5,198,346, and 5,096,815).

Alternatively, other forms of mutagenesis can be utilized to generate a combinatorial library. For example, AVMSP homologs (both agonist and antagonist forms) can be generated and isolated from a library by screening using, for example, alanine scanning mutagenesis and the like (Ruf et al., (1994) *Biochemistry* 33:1565-1572; Wang et al., (1994) *J. Biol. Chem.* 269:3095-3099; Balint et al., (1993) *Gene* 137:109-118; Grodberg et al., (1993) *Eur. J. Biochem.* 218:597-601; Nagashima et al., (1993) *J. Biol. Chem.* 268:2888-2892; Lowman et al., (1991) *Biochemistry* 30:10832-10838; and Cunningham et al., (1989) *Science* 244:1081-1085), by linker scanning mutagenesis (Gustin et al., (1993) *Virology* 193:653-660; Brown et al., (1992) *Mol. Cell Biol.* 12:2644-2652; McKnight et al., (1982) *Science* 232:316); by saturation mutagenesis (Meyers et al., (1986) *Science* 232:613); by PCR mutagenesis (Leung et al., (1989) *Method Cell Mol Biol* 1:11-19); or by random mutagenesis, including chemical mutagenesis, etc. (Miller et al., (1992) *A Short Course in Bacterial Genetics*, CSHL Press, Cold Spring Harbor, NY; and Greener et al., (1994)

Strategies in Mol Biol 7:32-34). Linker scanning mutagenesis, particularly in a combinatorial setting, is an attractive method for identifying truncated (bioactive) forms of AVMSPs.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations and truncations, and, for that matter, for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of AVMSP homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate sequences created by combinatorial mutagenesis techniques.

In an illustrative embodiment of a screening assay, candidate combinatorial gene products of one of the subject proteins are displayed on the surface of a cell or virus, and the ability of particular cells or viral particles to bind an AVMSP, eg. a protein designated in Tables 2 and 3, is detected in a "panning assay". For instance, a library of AVMSP-IP variants can be cloned into the gene for a surface membrane protein of a bacterial cell (Ladner et al., WO 88/06630; Fuchs et al., (1991) Bio/Technology 9:1370-1371; and Goward et al., (1992) TIBS 18:136-140), and the resulting fusion protein detected by panning, e.g. using a fluorescently labeled molecule which binds the AVMSP-IP, such as FITC-labelled AVMSP, to score for potentially functional homologs. Cells can be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, separated by a fluorescence-activated cell sorter.

In similar fashion, the gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at very high concentrations, a large number of phage can be screened at one time. Second, since each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical E. coli filamentous phages M13, fd, and fl are most often used in phage display libraries, as either of the phage gIII or gVIII coat

proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al., PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al., (1992) J. Biol. Chem. 267:16007-16010; Griffiths et al., (1993) EMBO J. 12:725-734; Clackson et al., (1991) Nature 352:624-628; and Barbas et al., (1992) PNAS USA 89:4457-4461).

The invention also provides for reduction of the subject AVMSPs to generate mimetics, e.g. peptide or non-peptide agents, which are able to mimic binding of the authentic protein to another cellular partner. Such mutagenic techniques as described above, as well as the thioredoxin system, are also particularly useful for mapping the determinants of an AVMSP which participate in protein-protein interactions involved in, for example, binding of proteins involved in viral maturation to each other. To illustrate, the critical residues of an AVMSP which are involved in molecular recognition of a substrate protein can be determined and used to generate AVMSP-derived peptidomimetics which bind to the substrate protein, and by inhibiting AVMSP binding, act to inhibit its biological activity. By employing, for example, scanning mutagenesis to map the amino acid residues of an AVMSP which are involved in binding to another polypeptide, peptidomimetic compounds can be generated which mimic those residues involved in binding. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman 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), 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, IL, 1985), b-turn dipeptide cores (Nagai et al., (1985) Tetrahedron Lett 26:647; and Sato et al., (1986) J Chem Soc Perkin Trans 1:1231), and b-aminoalcohols (Gordon et al., (1985) Biochem Biophys Res Commun 126:419; and Dann et al., (1986) Biochem Biophys Res Commun 134:71).

5. Antibodies and Uses Therefor

Another aspect of the invention pertains to an antibody specifically reactive with an AVMSP, e.g., a wild-type or mutated AVMSP. For example, by using immunogens derived from an AVMSP, e.g., based on the cDNA sequences, anti-protein/anti-peptide antisera or

monoclonal antibodies can be made by 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 the peptide (e.g., a mammalian AVMSP or an antigenic fragment which is capable of eliciting an antibody response, or a fusion protein as described above). 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 an AVMSP 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. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of an AVMSP of a mammal, e.g., antigenic determinants of a protein set forth in SEQ ID Nos: 1-39.

In one embodiment, antibodies are specific for a RING domain, an SH3 domain, an SH2 domain, and a C2 domain, and preferably the domain is part of an AVMSP, such as a domain of an AVMSP shown in one of SEQ ID NOs: 1-39. In another embodiment, the antibodies are immunoreactive with one or more proteins having an amino acid sequence that is at least 80% identical to an amino acid sequence as set forth in SEQ ID Nos. 1-39. In other embodiments, an antibody is immunoreactive with one or more proteins having an amino acid sequence that is 85%, 90%, 95%, 98%, 99% or identical to an amino acid sequence as set forth in SEQ ID Nos. 1-39.

In a further embodiment, an antibody of the invention disrupts the direct or indirect interaction between an AVMSP polypeptide and an AVMSP-AP.

Following immunization of an animal with an antigenic preparation of an AVMSP, anti- AVMSP antisera can be obtained and, if desired, polyclonal anti- AVMSP antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256: 495-497), the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4: 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a mammalian AVMSP polypeptide of the present invention and monoclonal antibodies isolated

from a culture comprising such hybridoma cells. In one embodiment anti-human AVMSF antibodies specifically react with the protein encoded by a nucleic acid having SEQ ID Nos 40-79.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject mammalian AVMSF 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 antibody of the present invention is further intended to include bispecific, single-chain, and chimeric and humanized molecules having affinity for an AVMSF protein conferred by at least one CDR region of the antibody. In preferred embodiments, the antibodies, the antibody further comprises a label attached thereto and able to be detected, (e.g., the label can be a radioisotope, fluorescent compound, enzyme or enzyme co-factor).

Anti-AVMSF antibodies can be used, e.g., to monitor AVMSF levels in an individual, particularly the presence of AVMSFs in the plasma membrane for determining whether or not said patient is infected with a virus such as an RNA virus, or allowing determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. In addition, AVMSFs are understood to localize, occasionally, to the released viral particle. Viral particles may be collected and assayed for the presence of one or more AVMSFs. The level of AVMSF may be measured in a variety of sample types such as, for example, cells and/or in bodily fluid, such as in blood samples.

Another application of anti-AVMSF antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors such as gt11, gt18-23, ZAP, and ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, gt11 will produce fusion proteins whose amino termini consist of β -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of an AVMSF, e.g., other orthologs of a particular protein or other paralogs from the same species, can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with the appropriate anti-AVMSF antibodies. Positive phage detected by this assay can then be isolated from the infected plate. Thus, the presence

of AVMSP homologs can be detected and cloned from other animals, as can alternate isoforms (including splice variants) from humans.

6. Homology Searching of Nucleotide and Polypeptide Sequences

The nucleotide or amino acid sequences of the invention may be used as query sequences against databases such as GenBank, SwissProt, BLOCKS, and Pima II. These databases contain previously identified and annotated sequences that can be searched for regions of homology (similarity) using BLAST, which stands for Basic Local Alignment Search Tool (Altschul S F (1993) J Mol Evol 36:290-300; Altschul, S F et al (1990) J Mol Biol 215:403-10).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith, R. F. and T. F. Smith (1992; Protein Engineering 5:35-51), incorporated herein by reference, can be used when dealing with primary sequence patterns and secondary structure gap penalties. As disclosed in this application, sequences have lengths of at least 49 nucleotides and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin and Altschul (1993; Proc Nat Acad Sci 90:5873-7) and incorporated herein by reference, searches matches between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. Preferably the threshold is set at 10-25 for nucleotides and 3-15 for peptides.

7. Diagnostic Assays

A further aspect of the invention includes diagnostic assays for determining whether a cell is infected with a virus and for characterizing the nature, progression and/or infectivity of the infection.

In one embodiment, it is contemplated that AVMSPs certain associated proteins localize to different regions of the cell depending on the function being performed. In the course of normal activities, it is expected that AVMSPs will be free in the cytoplasm or associated with an intracellular organelle, such as the nucleus, the Golgi network, etc. During a viral infection, certain AVMSPs are recruited to the cell membrane to participate in viral maturation, including ubiquitination and membrane fusion. As a result, the detection of an

AVMSP associated with the plasma membrane fraction is indicative of a viral infection. Additionally, the presence of AVMSPs at the plasma membrane would suggest that the infective virus is in the process of reproducing and is therefore actively engaged in infective or lytic activity (versus a lysogenic or otherwise dormant state).

Association of the proteins of the invention with the plasma membrane may be detected using a variety of techniques known in the art. For example, membrane preparations may be prepared by breaking open the cells (*via* sonication or detergent lysis) and then separating the membrane components from the cytosolic fraction *via* centrifugation. Segregation of proteins into the membrane fraction can be detected with antibodies specific for the protein of interest, for example by using Western blot analysis or ELISA techniques. Plasma membranes may be separated from intracellular membranes on the basis of density using density gradient centrifugation. Alternatively, plasma membranes may be obtained by chemically or enzymatically modifying the surface of the cell and affinity purifying the plasma membrane by selectively binding the modifications. An exemplary modification includes non-specific biotinylation of proteins at the cell surface. Plasma membranes may also be selected for by affinity purifying for abundant plasma membrane proteins.

Transmembrane AVMSP proteins containing an extracellular domain can be detected using FACS analysis. For FACS analysis, whole cells are incubated with a fluorescently labeled antibody (e.g., an FITC-labelled antibody) capable of recognizing the extracellular domain of the protein of interest. The level of fluorescent staining of the cells may then be determined by FACS analyses (see e.g., Weiss and Stobo, (1984) J. Exp. Med., 160:1284-1299). Such proteins are expected to reside on intracellular membranes in uninfected cells and the plasma membrane in infected cells. FACS analysis would fail to detect an extracellular domain unless the protein is present at the plasma membrane.

In a further embodiment, proteins associated with the membranes of cells and/or viral particles may be profiled. Profiling involves identifying the presence or absence of more than one protein in the membrane associated fraction of a sample. For example, the presence of AVMSPs are detected in the membrane associated fraction of cells obtained from a person suspected of a viral infection. Similar profiles may be developed from subjects infected by known viruses or subjects thought to be free of infection. Profiles may be compared to identify proteins that change in abundance, or qualitatively (eg. in terms of pI, molecular weight, or other indicators of post-translational modification). Profiles may be compiled into a database for computer-assisted comparisons. Comparison of profiles may be used to identify an AVMSP that is altered in response to a certain viral infection. This AVMSP may

then be used as a diagnostic for that type of viral infection. The AVMSP may also then be used as a target to identify therapeutic agents that will interfere with its function in the infection. Exemplary profiles of the invention will include information about the abundance of more than one AVMSP selected from those represented by SEQ ID Nos. 1-39. Other exemplary profiles will include information about the abundance of 5, 10, 20, 30, 40, 50, 60 or all of the proteins represented by SEQ ID Nos. 1-39.

Localization of the proteins of the invention may also be determined using histochemical techniques. For example, cells may be fixed and stained with a fluorescently labeled antibody specific for the protein of interest. The stained cells may then be examined under the microscope to determine the subcellular localization of the antibody bound proteins.

In addition, as noted above, AVMSPs may localize to released or budding viral particles. The presence of these proteins in viral particles may be determined by a variety of methods. For example, viral particles may be enriched and analyzed by Western blot or ELISA. As another example, viral particles or cells having budding viroids may be examined by electron microscopy. Immunogold labeling, for example, is useful for localizing AVMSPs by electron microscopy.

Samples to be used for diagnostic assays may include essentially any sample comprising cells and/or viral particles or a sample prepared from a cellular sample. Exemplary samples would include fluid samples (eg. blood, urine, saliva, mucus, bronchoalveolar lavage, cerebrospinal fluid etc.). Other fluids comprising cells and/or viral particles are well known to those of skill in the art. Other sample types include stool samples, tissue biopsies and any processed or purified form of the above.

8. Drug Screening Assays

The present invention also provides assays for identifying therapeutics which either interfere with or promote viral maturation, particularly by affecting AVMSP function. In one embodiment, the assay detects agents which inhibit interaction of one or more subject AVMSPs with an AVMSP-AP. In another embodiment, the assay detects agents which modulate the intrinsic biological activity of an AVMSP, AVMSP complex, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, and the like. Such modulators can be used, for example, in the treatment of viral infections and/or particularly viral infections by a virus that uses a Gag-dependent maturation system (eg. retrovirus, rhabdovirus, filovirus).

In one aspect, the invention provides methods and compositions for the identification of compositions that interfere with the function of AVMSPs. Given the critical role of AVMSPs in virion release, compositions that perturb the formation or stability of the protein-protein interactions between AVMSPs and the proteins that they interact with, such as AVMSP-APs, are candidate pharmaceuticals for the treatment of viral infections.

While not wishing to be bound to mechanism, it is postulated that AVMSPs promote the assembly of protein complexes that are critically important in release of virions. Complexes of the invention may include a combination of at least one of the following: a polypeptide comprising an AVMSP, a Gag protein, a Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIP, a GTPase, an E2 enzyme, tsg101, a cullin, AP-1, AP-2, and a clathrin.

The type of complex formed by an AVMSP will depend upon the domains present in the protein. While not intended to be limiting, exemplary domains of potential interacting proteins are provided below. An SH3 domain is expected to interact with one or more small GTPases, such as members of the Arf, Rab, Rac and Rho families. In addition, the following Table provides a list of exemplary protein domains that are associated with the formation of AVMSP complexes:

Table 5

	Domain Name	Interacting motif	Description
1	SH2	Yxx ϕ	The Src homology 2 (SH2) domain is a protein domain of about 100 amino-acid residues first identified as a conserved sequence region between the oncoproteins Src and Fps. Similar sequences were later found in many other intracellular signal-transducing proteins. SH2 domains function as regulatory modules of intracellular signalling cascades by interacting with high affinity to phosphotyrosine-containing target peptides in a sequence-specific and strictly phosphorylation-dependent manner.
2	SH3	PxRPxR (proline rich)	The Src homology 3 (SH3) domain is a small protein domain of about 60 amino-acid residues first identified as a conserved sequence in the non-catalytic part of several cytoplasmic protein tyrosine kinases (e.g. Src, Abl, Lck). Since then, it has been found in a great variety of other intracellular or membrane-associated proteins. The function of the SH3 domain is to mediate assembly of specific protein complexes via binding to proline-rich peptides.
3	Endocytosis motifs: μ 2 subunit of AP2 Badaptin or	Yxx ϕ (D/E)xxxLL (dileucine) FYRAL	Tyrosine not phosphorylated

	through an adaptor to AP2 (NEF) AP2? Clatherin	NPXY	
4	Clatherin assembly domain	Multiple DLL or SLL	AP-2, AP180, AP1
5	C2	Ca ²⁺ , phospholipids, inositol polyphosphates.	Ca ²⁺ -binding motif present in phospholipases, protein kinases C, and synaptotamins (among others). Some do not appear to contain Ca ²⁺ -binding sites. Particular C2s appear to bind phospholipids, inositol polyphosphates, and intracellular proteins. Unusual occurrence in perforin. Synaptotagmin and PLC C2s are permuted in sequence with respect to N- and C-terminal beta strands.
6	WW	[AP]-P-P-[AP]-Y	Found in dystrophin. The domain, which spans about 35 residues, is repeated up to 4 times in some proteins. It has been shown to bind proteins with particular proline-motifs, and thus resembles somewhat SH3 domains. The name WW or WWP derives from the presence of Trp as well as that of a conserved Pro. It is frequently associated with other domains typical for proteins in signal transduction processes.
7	RCC1	Ran GTPase	The regulator of chromosome condensation (RCC1) is a eukaryotic protein, which binds to chromatin and interacts with ran, a nuclear GTP-binding protein to promote the loss of bound GDP and the uptake of fresh GTP, thus acting as a guanine-nucleotide dissociation stimulator (GDS).

A variety of assay formats will suffice and, in light of the present disclosure, those not expressly described herein will nevertheless be comprehended by one of ordinary skill in the art. Assay formats which approximate such conditions as formation of protein complexes, enzymatic activity, and even an AVMSF-mediated membrane reorganization activity, can be generated in many different forms, and include assays based on cell-free systems, e.g. purified proteins or cell lysates, as well as cell-based assays which utilize intact cells. Simple binding assays can also be used to detect agents which, by disrupting the binding of AVMSFs to interacting protein, or the binding of an AVMSF or complex to a substrate, can inhibit viral maturation. Agents to be tested for their ability to act as viral maturation inhibitors can be produced, for example, by bacteria, yeast or other organisms (e.g. natural products), produced chemically (e.g. small molecules, including peptidomimetics), or produced recombinantly. In a preferred embodiment, the test agent is a small organic molecule, e.g., other than a peptide or oligonucleotide, having a molecular weight of less than about 2,000 daltons.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays of the present invention which are performed in

cell-free systems, such as may be developed with purified or semi-purified proteins or with lysates, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or changes in enzymatic properties of the molecular target.

In preferred in vitro embodiments of the present assay, a reconstituted AVMS complex comprises a reconstituted mixture of at least semi-purified proteins. By semi-purified, it is meant that the proteins utilized in the reconstituted mixture have been previously separated from other cellular or viral proteins. For instance, in contrast to cell lysates, the proteins involved in AVMS complex formation, are present in the mixture to at least 50% purity relative to all other proteins in the mixture, and more preferably are present at 90-95% purity. In certain embodiments of the subject method, the reconstituted protein mixture is derived by mixing highly purified proteins such that the reconstituted mixture substantially lacks other proteins (such as of cellular or viral origin) which might interfere with or otherwise alter the ability to measure AVMS complex assembly and/or disassembly.

Assaying AVMS complexes, in the presence and absence of a candidate inhibitor, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes.

In one embodiment of the present invention, drug screening assays can be generated which detect inhibitory agents on the basis of their ability to interfere with assembly or stability of the AVMS complex. In an exemplary binding assay, the compound of interest is contacted with a mixture comprising an AVMS polypeptide and at least one interacting polypeptide. Detection and quantification of AVMS complexes provides a means for determining the compound's efficacy at inhibiting (or potentiating) interaction between the two polypeptides. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, the formation of complexes is quantitated in the absence of the test compound.

Complex formation between the AVMS polypeptides or between an AVMS and a substrate polypeptide may be detected by a variety of techniques, many of which are effectively described above. For instance, modulation in the formation of complexes can be

quantitated using, for example, detectably labeled proteins (e.g. radiolabeled, fluorescently labeled, or enzymatically labeled), by immunoassay, or by chromatographic detection. Surface plasmon resonance systems, such as those available from BioCore, Inc., may also be used to detect protein-protein interaction

Often, it will be desirable to immobilize one of the polypeptides to facilitate separation of complexes from uncomplexed forms of one of the proteins, as well as to accommodate automation of the assay. In an illustrative embodiment, a fusion protein can be provided which adds a domain that permits the protein to be bound to an insoluble matrix. For example, GST-AVMSP or -AMVSP fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with a potential interacting protein, e.g. an ³⁵S-labeled polypeptide, and the test compound and incubated under conditions conducive to complex formation. Following incubation, the beads are washed to remove any unbound interacting protein, and the matrix bead-bound radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after the complexes are dissociated, e.g. when microtitre plate is used. Alternatively, after washing away unbound protein, the complexes can be dissociated from the matrix, separated by SDS-PAGE gel, and the level of interacting polypeptide found in the matrix-bound fraction quantitated from the gel using standard electrophoretic techniques.

In yet another embodiment, the AVMSP and potential interacting polypeptide can be used to generate an interaction trap assay (see also, U.S. Patent No: 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J Biol Chem 268:12046-12054; Bartel et al. (1993) Biotechniques 14:920-924; and Iwabuchi et al. (1993) Oncogene 8:1693-1696), for subsequently detecting agents which disrupt binding of the proteins to one and other.

In particular, the method makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a DNA-binding domain of a transcriptional activator can be fused in frame to the coding sequence for a "bait" protein, e.g., an AVMSP polypeptide of sufficient length to bind to a potential interacting protein. The second hybrid protein encodes a transcriptional activation domain fused in frame to a gene encoding a "fish" protein, e.g., a potential interacting protein of sufficient length to interact with the AVMSP polypeptide portion of the bait fusion protein. If the bait and fish proteins are able to interact, e.g., form an AVMSP complex, they bring into close proximity the two domains of the transcriptional activator. This proximity causes transcription of a reporter gene which is operably linked to a transcriptional regulatory site responsive to the

transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of the bait and fish proteins.

In accordance with the present invention, the method includes providing a host cell, preferably a yeast cell, e.g., *Kluyverei lactis*, *Schizosaccharomyces pombe*, *Ustilago maydis*, *Saccharomyces cerevisiae*, *Neurospora crassa*, *Aspergillus niger*, *Aspergillus nidulans*, *Pichia pastoris*, *Candida tropicalis*, and *Hansenula polymorpha*, though most preferably *S. cerevisiae* or *S. pombe*. The host cell contains a reporter gene having a binding site for the DNA-binding domain of a transcriptional activator used in the bait protein, such that the reporter gene expresses a detectable gene product when the gene is transcriptionally activated. The first chimeric gene may be present in a chromosome of the host cell, or as part of an expression vector. Interaction trap assays may also be performed in mammalian and bacterial cell types.

The host cell also contains a first chimeric gene which is capable of being expressed in the host cell. The gene encodes a chimeric protein, which comprises (i) a DNA-binding domain that recognizes the responsive element on the reporter gene in the host cell, and (ii) a bait protein, such as an AVMSPP polypeptide sequence.

A second chimeric gene is also provided which is capable of being expressed in the host cell, and encodes the "fish" fusion protein. In one embodiment, both the first and the second chimeric genes are introduced into the host cell in the form of plasmids. Preferably, however, the first chimeric gene is present in a chromosome of the host cell and the second chimeric gene is introduced into the host cell as part of a plasmid.

Preferably, the DNA-binding domain of the first hybrid protein and the transcriptional activation domain of the second hybrid protein are derived from transcriptional activators having separable DNA-binding and transcriptional activation domains. For instance, these separate DNA-binding and transcriptional activation domains are known to be found in the yeast GAL4 protein, and are known to be found in the yeast GCN4 and ADR1 proteins. Many other proteins involved in transcription also have separable binding and transcriptional activation domains which make them useful for the present invention, and include, for example, the LexA and VP16 proteins. It will be understood that other (substantially) transcriptionally-inert DNA-binding domains may be used in the subject constructs; such as domains of ACE1, IclI, lac repressor, jun or fos. In another embodiment, the DNA-binding domain and the transcriptional activation domain may be from different proteins. The use of a LexA DNA binding domain provides certain advantages. For example, in yeast, the LexA moiety contains no activation function and has no known effect on transcription of yeast

genes. In addition, use of LexA allows control over the sensitivity of the assay to the level of interaction (see, for example, the Brent et al. PCT publication WO94/10300).

In preferred embodiments, any enzymatic activity associated with the bait or fish proteins is inactivated, e.g., dominant negative or other mutants of an AVMSF can be used.

Continuing with the illustrated example, the AVMSF-mediated interaction, if any, between the bait and fish fusion proteins in the host cell, therefore, causes the activation domain to activate transcription of the reporter gene. The method is carried out by introducing the first chimeric gene and the second chimeric gene into the host cell, and subjecting that cell to conditions under which the bait and fish fusion proteins are expressed in sufficient quantity for the reporter gene to be activated. The formation of an AVMSF/interacting protein complex results in a detectable signal produced by the expression of the reporter gene. Accordingly, the level of formation of a complex in the presence of a test compound and in the absence of the test compound can be evaluated by detecting the level of expression of the reporter gene in each case. Various reporter constructs may be used in accord with the methods of the invention and include, for example, reporter genes which produce such detectable signals as selected from the group consisting of an enzymatic signal, a fluorescent signal, a phosphorescent signal and drug resistance.

One aspect of the present invention provides reconstituted protein preparations including an AVMSF and one or more interacting polypeptides.

In still further embodiments of the present assay, the AVMSF complex is generated in whole cells, taking advantage of cell culture techniques to support the subject assay. For example, as described below, the AVMSF complex can be constituted in a eukaryotic cell culture system, including mammalian and yeast cells. Often it will be desirable to express one or more viral proteins (eg. Gag or Env) in such a cell along with a subject AVMSF. It may also be desirable to infect the cell with a virus of interest. Advantages to generating the subject assay in an intact cell include the ability to detect inhibitors which are functional in an environment more closely approximating that which therapeutic use of the inhibitor would require, including the ability of the agent to gain entry into the cell. Furthermore, certain of the in vivo embodiments of the assay, such as examples given below, are amenable to high through-put analysis of candidate agents.

The components of the AVMSF can be endogenous to the cell selected to support the assay. Alternatively, some or all of the components can be derived from exogenous sources. For instance, fusion proteins can be introduced into the cell by recombinant techniques (such

as through the use of an expression vector), as well as by microinjecting the fusion protein itself or mRNA encoding the fusion protein.

In any case, the cell is ultimately manipulated after incubation with a candidate drug and assayed for an AVMSF activity. AVMSF activities may include, without limitation, complex formation, ubiquitination and membrane fusion events (eg. release of viral buds or fusion of vesicles). AVMSF complex formation may be assessed by immunoprecipitation and analysis of co-immunoprecipitated proteins or affinity purification and analysis of co-purified proteins. Fluorescence Resonance Energy Transfer (FRET)-based assays may also be used to determine complex formation. Fluorescent molecules having the proper emission and excitation spectra that are brought into close proximity with one another can exhibit FRET. The fluorescent molecules are chosen such that the emission spectrum of one of the molecules (the donor molecule) overlaps with the excitation spectrum of the other molecule (the acceptor molecule). The donor molecule is excited by light of appropriate intensity within the donor's excitation spectrum. The donor then emits the absorbed energy as fluorescent light. The fluorescent energy it produces is quenched by the acceptor molecule. FRET can be manifested as a reduction in the intensity of the fluorescent signal from the donor, reduction in the lifetime of its excited state, and/or re-emission of fluorescent light at the longer wavelengths (lower energies) characteristic of the acceptor. When the fluorescent proteins physically separate, FRET effects are diminished or eliminated. (U.S. Patent No. 5,981,200).

For example, a cyan fluorescent protein is excited by light at roughly 425 - 450 nm wavelength and emits light in the range of 450 - 500 nm. Yellow fluorescent protein is excited by light at roughly 500 - 525 nm and emits light at 525 - 500 nm. If these two proteins are placed in solution, the cyan and yellow fluorescence may be separately visualized. However, if these two proteins are forced into close proximity with each other, the fluorescent properties will be altered by FRET. The bluish light emitted by CFP will be absorbed by YFP and re-emitted as yellow light. This means that when the proteins are stimulated with light at wavelength 450 nm, the cyan emitted light is greatly reduced and the yellow light, which is not normally stimulated at this wavelength, is greatly increased. FRET is typically monitored by measuring the spectrum of emitted light in response to stimulation with light in the excitation range of the donor and calculating a ratio between the donor-emitted light and the acceptor-emitted light. When the donor:acceptor emission ratio is high, FRET is not occurring and the two fluorescent proteins are not in close proximity. When the donor: acceptor emission ratio is low, FRET is occurring and the two fluorescent proteins are

in close proximity. In this manner, the interaction between a first and second polypeptide may be measured.

The occurrence of FRET also causes the fluorescence lifetime of the donor fluorescent moiety to decrease. This change in fluorescence lifetime can be measured using a technique termed fluorescence lifetime imaging technology (FLIM) (Verveer et al. (2000) *Science* 290: 1567-1570; Squire et al. (1999) *J. Microsc.* 193: 36; Verveer et al. (2000) *Biophys. J.* 78: 2127). Global analysis techniques for analyzing FLIM data have been developed. These algorithms use the understanding that the donor fluorescent moiety exists in only a limited number of states each with a distinct fluorescence lifetime. Quantitative maps of each state can be generated on a pixel-by-pixel basis.

To perform FRET-based assays, the AVMSF and the interacting protein of interest are both fluorescently labeled. Suitable fluorescent labels are, in view of this specification, well known in the art. Examples are provided below, but suitable fluorescent labels not specifically discussed are also available to those of skill in the art. Fluorescent labeling may be accomplished by expressing a polypeptide as a fusion protein with a fluorescent protein, for example fluorescent proteins isolated from jellyfish, corals and other coelenterates. Exemplary fluorescent proteins include the many variants of the green fluorescent protein (GFP) of *Aequoria victoria*. Variants may be brighter, dimmer, or have different excitation and/or emission spectra. Certain variants are altered such that they no longer appear green, and may appear blue, cyan, yellow or red (termed BFP, CFP, YFP and RFP, respectively). Fluorescent proteins may be stably attached to polypeptides through a variety of covalent and noncovalent linkages, including, for example, peptide bonds (eg. expression as a fusion protein), chemical cross-linking and biotin-streptavidin coupling. For examples of fluorescent proteins, see U.S. Patents 5,625,048; 5,777,079; 6,066,476; 6,124,128; Prasher et al. (1992) *Gene*, 111:229-233; Heim et al. (1994) *Proc. Natl. Acad. Sci., USA*, 91:12501-04; Ward et al. (1982) *Photochem. Photobiol.*, 35:803-808 ; Levine et al. (1982) *Comp. Biochem. Physiol.*, 72B:77-85; Tersikh et al. (2000) *Science* 290: 1585-88.

Other exemplary fluorescent moieties well known in the art include derivatives of fluorescein, benzoxadioazole, coumarin, eosin, Lucifer Yellow, pyridyloxazole and rhodamine. These and many other exemplary fluorescent moieties may be found in the *Handbook of Fluorescent Probes and Research Chemicals* (2000, Molecular Probes, Inc.), along with methodologies for modifying polypeptides with such moieties. Exemplary proteins that fluoresce when combined with a fluorescent moiety include, yellow fluorescent protein from *Vibrio fischeri* (Baldwin et al. (1990) *Biochemistry* 29:5509-15), peridinin-

chlorophyll a binding protein from the dinoflagellate *Symbiodinium sp.* (Morris et al. (1994) *Plant Molecular Biology* 24:673:77) and phycobiliproteins from marine cyanobacteria such as *Synechococcus*, e.g., phycoerythrin and phycocyanin (Wilbanks et al. (1993) *J. Biol. Chem.* 268:1226-35). These proteins require flavins, peridinin-chlorophyll a and various phycobilins, respectively, as fluorescent co-factors.

FRET-based assays may be used in cell-based assays and in cell-free assays. FRET-based assays are amenable to high-throughput screening methods including Fluorescence Activated Cell Sorting and fluorescent scanning of microtiter arrays.

10. Methods and Compositions for Treatment of Viral Disorders

In a further aspect, the invention provides methods and compositions for treatment of viral disorders, and particularly disorders caused by RNA viruses, including but not limited to retroviruses, rhabdoviruses and filoviruses. Preferred therapeutics of the invention function by disrupting the biological activity of an AVMSF or AVMSF complex in viral maturation.

Exemplary therapeutics of the invention include antisense therapies, RNAi or siRNA therapies, polypeptides, peptidomimetics, antibodies and small molecules.

Antisense therapies of the invention include methods of introducing antisense nucleic acids to disrupt the expression of AVMSFs or proteins that are necessary for AVMSF function. RNAi or siRNA therapies both refer to the administration of short double stranded RNA molecules that are complementary to a portion of an AVMSF transcript to achieve a decrease in the expression of that AVMSF and, thereby, a decrease in the function of that AVMSF. RNA molecules for siRNA may be prepared as a therapeutic composition by, for example, mixing them with a suitable carrier, such as a lipid or polycationic carrier that facilitates delivery of the nucleic acid across the membrane of the targeted cells.

Therapeutic polypeptides may be generated by designing polypeptides to mimic certain protein domains important in the formation of AVMSF complexes. For example, a polypeptide comprising an SH3 domain will compete for binding to an SH3 domain and will therefore act to disrupt binding of a Gag protein, for example, to the AVMSF complex. Likewise, a polypeptide that resembles an L domain may disrupt recruitment of Gag to the AVMSF complex. Such polypeptide mimetics may be targeted to any of a variety of domains, including for example, those domains listed in Table 5.

In view of the specification, methods for generating antibodies directed to epitopes of AVMSFs and AVMSF-interacting proteins are known in the art. Antibodies may be introduced into cells by a variety of methods. One exemplary method comprises generating a

nucleic acid encoding a single chain antibody that is capable of disrupting an AVMSP complex. Such a nucleic acid may be conjugated to antibody that binds to receptors on the surface of target cells. It is contemplated that in certain embodiments, the antibody may target viral proteins that are present on the surface of infected cells, and in this way deliver the nucleic acid only to infected cells. Once bound to the target cell surface, the antibody is taken up by endocytosis, and the conjugated nucleic acid is transcribed and translated to produce a single chain antibody that interacts with and disrupts the targeted AVMSP complex. Nucleic acids expressing the desired single chain antibody may also be introduced into cells using a variety of more conventional techniques, such as viral transfection (eg. using an adenoviral system) or liposome-mediated transfection.

Small molecules of the invention may be identified for their ability to modulate the formation of AVMSP complexes, as described above.

In view of the teachings herein, one of skill in the art will understand that the methods and compositions of the invention are applicable to a wide range of RNA viruses including retroviruses. While not intended to be limiting, relevant retroviruses include: C-type retrovirus which causes lymphosarcoma in Northern Pike, the C-type retrovirus which infects mink, the caprine lentivirus which infects sheep, the Equine Infectious Anemia Virus (EIAV), the C-type retrovirus which infects pigs, the Avian Leukosis Sarcoma Virus (ALSV), the Feline Leukemia Virus (FeLV), the Feline Aids Virus, the Bovine Leukemia Virus (BLV), the Simian Leukemia Virus (SLV), the Simian Immuno-deficiency Virus (SIV), the Human T-cell Leukemia Virus type-I (HTLV-I), the Human T-cell Leukemia Virus type-II (HTLV-II), Human Immunodeficiency virus type-2 (HIV-2) and Human Immunodeficiency virus type-1 (HIV-1). Other RNA viruses include picornaviruses such as enterovirus, poliovirus, coxsackievirus and hepatitis A virus, the caliciviruses, including Norwalk-like viruses, the rhabdoviruses, including rabies virus, the togaviruses including alphaviruses, Semliki Forest virus, denguevirus, yellow fever virus and rubella virus, the orthomyxoviruses, including Type A, B, and C influenza viruses, the bunyaviruses, including the Rift Valley fever virus and the hantavirus, the filoviruses such as Ebola virus and Marburg virus, and the paramyxoviruses, including mumps virus and measles virus.

11. Effective Dose

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for

determining The Ld50 (The Dose Lethal To 50% Of The Population) And The Ed50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic induces are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

12. Formulation and Use

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by, for example, injection, inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For such therapy, the compounds of the invention can be formulated for a variety of loads of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous. For injection, the compounds of the invention can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition,

the compounds may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., ationd oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound. For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or

dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art. A wash solution can be used locally to treat an injury or inflammation to accelerate healing.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

For therapies involving the administration of nucleic acids, the oligomers of the invention can be formulated for a variety of modes of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, intranodal, and subcutaneous for injection, the oligomers of the invention can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the oligomers may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

Systemic administration can also be by transmucosal or transdermal means, or the compounds can be administered orally. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For oral administration, the oligomers are formulated into conventional oral administration forms such as capsules, tablets, and tonics. For topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art.

EXAMPLES

1. Role of PRT3 in virus-like particle (VLP) budding

1. Objective:

Use RNAi to inhibit PRT3 gene expression and compare the efficiency of viral budding and Gag expression and processing in treated and untreated cells.

2. Study Plan:

HeLa SS-6 cells are transfected with mRNA-specific RNAi in order to knockdown the target proteins. Since maximal reduction of target protein by RNAi is achieved after 48 hours, cells are transfected twice – first to reduce target mRNAs, and subsequently to express the viral Gag protein. The second transfection is performed with pNLenv (plasmid that encodes HIV) and with low amounts of RNAi to maintain the knockdown of target protein during the time of gag expression and budding of VLPs. Reduction in mRNA levels due to RNAi effect is verified by RT-PCR amplification of target mRNA.

3. Methods, Materials, Solutions

a. Methods

- i. Transfections according to manufacturer's protocol and as described in procedure.
- ii. Protein determined by Bradford assay.
- iii. SDS-PAGE in Hoeffer miniVE electrophoresis system. Transfer in Bio-Rad mini-protean II wet transfer system. Blots visualized using Typhoon system, and ImageQuant software (ABbiotech)

b. Materials

Material	Manufacturer	Catalog #	Batch #
Lipofectamine 2000 (LF2000)	Life Technologies	11668-019	1112496
OptiMEM	Life Technologies	31985-047	3063119
RNAi Lamin A/C	Self	13	
RNAi TSG101 688	Self	65	
RNAi PRT3 524	Self	81	
plenvl1 PTAP	Self	148	
plenvl1 ATAP	Self	149	
Anti-p24 polyclonal antibody	Seramun		A-0236/5-10-01
Anti-Rabbit Cy5 conjugated antibody	Jackson	144-175-115	48715
10% acrylamide Tris-Glycine SDS-PAGE gel	Life Technologies	NP0321	1081371
Nitrocellulose membrane	Schleicher & Schuell	401353	BA-83
NuPAGE 20X transfer buffer	Life Technologies	NP0006-1	224365
0.45µm filter	Schleicher & Schuell	10462100	CS1018-1

c. Solutions

Lysis Buffer	Compound	Concentration
	Tris-HCl pH 7.6	50mM
	MgCl ₂	15mM
	NaCl	150mM
	Glycerol	10%
	EDTA	1mM
	EGTA	1mM

	ASB-14 (add immediately before use)	1%
6X Sample Buffer	Tris-HCl, pH=6.8	1M
	Glycerol	30%
	SDS	10%
	DTT	9.3%
	Bromophenol Blue	0.012%
TBS-T	Tris pH=7.6	20mM
	NaCl	137mM
	Tween-20	0.1%

4. Procedure

a. Schedule

Day				
1	2	3	4	5
Plate cells	Transfection I (RNAi only)	Passage cells (1:3)	Transfection II (RNAi and pNlenv) (12:00, PM)	Extract RNA for RT-PCR (post transfection)
			Extract RNA for RT-PCR (pre-transfection)	Harvest VLPs and cells

b. Day 1

Plate HeLa SS-6 cells in 6-well plates (35mm wells) at concentration of 5×10^5 cells/well.

c. Day 2

2 hours before transfection replace growth medium with 2 ml growth medium without antibiotics.

Transfection I:

Reaction	RNAi name	TAGDA#	Reactions	RNAi [nM]	RNAi [20μM]	A	B
						OPTiMEM	LF2000 mix

					μ l	(μ l)	(μ l)
1	Lamin A/C	13	2	50	12.5	500	500
2	Lamin A/C	13	1	50	6.25	250	250
3	TSG101 688	65	2	20	5	500	500
5	PRT3 524	81	2	50	12.5	500	500

Transfections:

Prepare LF2000 mix : 250 μ l OptiMEM + 5 μ l LF2000 for each reaction. Mix by inversion, 5 times. Incubate 5 minutes at room temperature.

Prepare RNA dilution in OptiMEM (Table 1, column A). Add LF2000 mix dropwise to diluted RNA (Table 1, column B). Mix by gentle vortex. Incubate at room temperature 25 minutes, covered with aluminum foil.

Add 500 μ l transfection mixture to cells dropwise and mix by rocking side to side.

Incubate overnight.

d. Day 3

Split 1:3 after 24 hours. (Plate 4 wells for each reaction, except reaction 2 which is plated into 3 wells.)

e. Day 4

2 hours pre-transfection replace medium with DMEM growth medium without antibiotics.

Transfection II

					A	B	C	D
					Plasmid	RNAi		
					for 2.4	[20 μ M] for 10nM	OPTiMEM	LF2000 mix
Reaction	RNAi name	TAGDA#	Plasmid	Reactions	μ g (μ g/ μ l)	μ g (μ l)	(μ l)	(μ l)
1	Lamin A/C	13	PTAP	3	3.4	3.75	750	750
2	Lamin A/C	13	ATAP	3	2.5	3.75	750	750
3	TSG101 688	65	PTAP	3	3.4	3.75	750	750
5	PRT3 524	81	PTAP	3	3.4	3.75	750	750

Prepare LF2000 mix: 250 μ l OptiMEM + 5 μ l LF2000 for each reaction. Mix by inversion, 5 times. Incubate 5 minutes at room temperature.

Prepare RNA+DNA diluted in OptiMEM (Transfection II, A+B+C)

Add LF2000 mix (Transfection II, D) to diluted RNA+DNA dropwise, mix by gentle vortex, and incubate 1h while protected from light with aluminum foil.

Add LF2000 and DNA+RNA to cells, 500 μ l/well, mix by gentle rocking and incubate overnight.

f. Day 5

Collect samples for VLP assay (approximately 24 hours post-transfection) by the following procedure (cells from one well from each sample is taken for RNA assay, by RT-PCR).

g. Cell Extracts

i. Pellet floating cells by centrifugation (5min, 3000rpm at 40°C), save supernatant (continue with supernatant immediately to step h), scrape remaining cells in the medium which remains in the well, add to the corresponding floating cell pellet and centrifuge for 5 minutes, 1800rpm at 40°C.

ii. Wash cell pellet twice with ice-cold PBS.

iii. Resuspend cell pellet in 100 μ l lysis buffer and incubate 20 minutes on ice.

iv. Centrifuge at 14,000rpm for 15min. Transfer supernatant to a clean tube. This is the cell extract.

v. Prepare 10 μ l of cell extract samples for SDS-PAGE by adding SDS-PAGE sample buffer to 1X, and boiling for 10 minutes. Remove an aliquot of the remaining sample for protein determination to verify total initial starting material. Save remaining cell extract at -80 °C.

h. Purification of VLPs from cell media

i. Filter the supernatant from step g through a 0.45m filter.

ii. Centrifuge supernatant at 14,000rpm at 40C for at least 2h.

iii. Aspirate supernatant carefully.

iv. Re-suspend VLP pellet in hot (100°C warmed for 10 min at least) 1X sample buffer.

v. Boil samples for 10 minutes, 100°C.

i. Western Blot analysis

i. Run all samples from stages A and B on Tris-Glycine SDS-PAGE 10% (120V for 1.5h.).

ii. Transfer samples to nitrocellulose membrane (65V for 1.5h.).

iii. Stain membrane with ponceau S solution.

- iv. Block with 10% low fat milk in TBS-T for 1h.
- v. Incubate with anti p24 rabbit 1:500 in TBS-T o/n.
- vi. Wash 3 times with TBS-T for 7min each wash.
- vii. Incubate with secondary antibody anti rabbit cy5 1:500 for 30min.
- viii. Wash five times for 10min in TBS-T
- ix. View in Typhoon gel imaging system (Molecular Dynamics/APBiotech) for fluorescence signal.

Results are shown in Figures 82-84.

2. Exemplary PRT3 RT-PCR primers and siRNA duplexes

RT-PCR primers

	Name	Position	Sequence
Sense primer	PRT3=271	271	5' CTTGCCTTGCCAGCATAC 3' (SEQ ID NO:12)
Anti-sense primer	PRT3=926c	926C	5' CTGCCAGCATTCTTCAG 3' (SEQ ID NO:13)

siRNA duplexes:

siRNA No: 153
 siRNA Name: PRT3-230
 Position in mRNA 426-446
 Target sequence: 5' AACAGAGGCCTTGGAACCTG 3' SEQ ID NO: 14
 siRNA sense strand: 5' dTdTcAGAGGCCUUGGAAACCUG 3' SEQ ID NO: 15
 siRNA anti-sense strand: 5'dTdTcAGGUUCCAAGGCCUCUG 3' SEQ ID NO: 16

siRNA No: 155
 siRNA Name: PRT3-442
 Position in mRNA 638-658
 Target sequence: 5' AAAGAGCCTGGAGACCTTAA 3' SEQ ID NO: 17
 siRNA sense strand: 5' ddTdTAgAGCCUGGAGACCUUAAA 3' SEQ ID NO: 18
 siRNA anti-sense strand: 5' ddTdTUUUAAGGUCUCCAGGCUCU 3' SEQ ID NO: 19

siRNA No: 157
 siRNA Name: PRT3-U111

Position in mRNA	2973-2993	
Target sequence:	5' AAGGATTGGTATGTGACTCTG 3'	SEQ ID NO: 20
siRNA sense strand:	5' dTdTGGAUUGGUAUGUGACUCUG 3'	SEQ ID NO: 21
siRNA anti-sense strand:	5' dTdTTCAGAGUCACAUACCAAUCC 3'	SEQ ID NO: 22
siRNA No:	159	
siRNA Name:	PRT3-U410	
Position in mRNA	3272-3292	
Target sequence:	5' AAGCTGGATTATCTCCTGTTG 3'	SEQ ID NO: 23
siRNA sense strand:	5' ddTdTGCUGGAUUAUCUCCUGUUG 3'	SEQ ID NO: 24
siRNA anti-sense strand:	5' ddTdTCAACAGGAGAUAAUCCAGC 3'	SEQ ID NO: 25

3. Effects of PRT3 RNAi on HIV Release: Kinetics

A1. Transfections

1. One day before transfection plate cells at a concentration of 5×10^6 cell/well in 15cm plates.
2. Two hours before transfection, replace cell media to 20ml complete DMEM without antibiotics.
3. DNA dilution: for each transfection dilute 62.5 μ l RNAi in 2.5ml OptiMEM according to the table below. RNAi stock is 20 μ M (recommended concentration: 50nM, dilution in total medium amount 1:400).
4. LF 2000 dilution: for each transfection dilute 50 μ l LF 2000 reagent in 2.5ml OptiMEM.
5. Incubate diluted RNAi and LF 2000 for 5 minutes at RT.
6. Mix the diluted RNAi with diluted LF2000 and incubated for 20-25 minutes at RT.
7. Add the mixture to the cells (drop wise) and incubate for 24 hours at 37°C in CO₂ incubator.
8. One day after RNAi transfection split cells (in complete MEM medium to 2 15cm plate and 1 well in a 6 wells plate)
9. One day after cell split perform HIV transfection.
10. 6 hours after HIV transfection replace medium to complete MEM medium.

* RT-PCR for PRT3 may be performed to assess the degree of knockdown.

A2. Total RNA purification.

1. One day after transfection, wash cells twice with sterile PBS. .
2. Scrape cells in 2.3ml/200 μ l (for 15cm plate/1 well of a 6 wells plate) Tri reagent (with sterile scrapers) and freeze in -70°C.

Treatment	Chase time (hours)	Fraction	Labeling
Control=WT	1	Cells	A1
		VLP	A1 V
	2	Cells	A2
		VLP	A2 V
	3	Cells	A3
		VLP	A3 V
	4	Cells	A4
		VLP	A4 V
	5	Cells	A5
		VLP	A5 V
PRT3+WT	1	Cells	B1
		VLP	B1 V
	2	Cells	B2
		VLP	B2 V
	3	Cells	B3
		VLP	B3 V
	4	Cells	B4
		VLP	B4 V
	5	Cells	B5
		VLP	B5 V

B. Labeling

1. Take out starvation medium, thaw and place at 37°C.
2. Scrape cells in growth medium and transfer gently into 15 ml conical tube.
3. Centrifuge to pellet cells at 1800rpm for 5 minutes at room temperature.
4. Aspirate supernatant and let tube stand for 10 sec. Remove the rest of the supernatant with a 200µl pipetman.
5. Gently add 10ml warm starvation medium and resuspend carefully with a 10ml pipette, up and down, just turning may not resolve the cell pellet).
6. Transfer cells to 10cm tube and place in the incubator for 60 minutes. Set an Eppendorf thermo mixer to 37°C.
7. Centrifuge to pellet cells at 1800rpm for 5 minutes at room temperature.
8. Aspirate supernatant and let tube stand for 10 sec. Remove the rest of the supernatant with a 200 µl pipetman.
9. Cut a 200µl tip from the end and resuspend cells ($\sim 1.5 \times 10^7$ cells in 150 µl RPIM without Met, but try not to go over 250 µl if you have more cells) gently in 150 µl starvation medium. Transfer cells to an Eppendorf tube and place in the thermo mixer. Wait 10 sec and transfer the rest of the cells from the 10 ml tube to the Eppendorf tube, if necessary add another 50 µl to splash the rest of the cells out (all specimens should have the same volume of labeling reaction).
10. Pulse: Add 50 µl of ^{35}S -methionine (specific activity 14.2 µCi/µl), tightly cup tubes and place in thermo mixer. Set the mixing speed to the lowest possible (700 rpm) and incubate for 25 minutes.
11. Stop the pulse by adding 1ml ice-cold chase/stop medium. Shake tube very gently three times and pellet cells at 6000rpm for 6 sec.
12. Remove supernatant with a 1ml tip. Add gently 1ml ice-cold chase/stop medium to the pelleted cells and invert gently to resuspend.
13. Chase: Transfer all tubes to the thermo mixer and incubate for the required chase time (**830**: 1,2,3,4 and 5 hours; **828**: 3 hours only). At the end of total chase time, place tubes on ice, add 1ml ice-cold chase/stop and pellet cells for 1 minute at 14,000 rpm. Remove supernatant and transfer supernatant to a second eppendorf tube. The cell pellet freeze at -80°C, until all tubes are ready.
14. Centrifuge supernatants for 2 hours at 14,000rpm, 4°C. Remove the supernatant very gently, leave 20 µl in the tube (labeled as V) and freeze at -80°C until the end of the time course.

*** All steps are done on ice with ice-cold buffers

15. When the time course is over, remove all tubes from -80°C . Lyse VLP pellet (from step 14) and cell pellet (step 13) by adding 500 μl of lysis buffer (see solutions), resuspend well by pipeting up and down three times. Incubate on ice for 15 minutes, and spin in an eppendorf centrifuge for 15 minutes at 4°C , 14,000 rpm. Remove supernatant to a fresh tube, discard pellet.
16. Perform IP with anti-p24 sheep for all samples.

C. Immunoprecipitation

1. Preclearing: add to all samples 15 μl ImmunoPure PlusG (Pierce). Rotate for 1 hour at 4°C in a cyclor, spin 5 min at 4°C , and transfer to a new tube for IP.
 2. Add to all samples 20 μl of p24-protein G conjugated beads and incubate 4 hours in a cyclor at 4°C .
 3. Post immunoprecipitations, transfer all immunoprecipitations to a fresh tube.
 4. Wash beads once with high salt buffer, once with medium salt buffer and once with low salt buffer. After each spin don't remove all solution, but leave 50 μl solution on the beads. After the last spin remove supernatant carefully with a loading tip and leave ~ 10 μl solution.
 5. Add to each tube 20 μl 2x SDS sample buffer. Heat to 70°C for 10 minutes.
 6. Samples were separated on 10% SDS-PAGE.
 7. Fix gel in 25% ethanol and 10% acetic acid for 15 minutes.
 8. Pour off the fixation solution and soak gels in Amplify solution (NAMP 100 Amersham) for 15 minutes.
 9. Dry gels on warm plate ($60-80^{\circ}\text{C}$) under vacuum.
 10. Expose gels to screen for 2 hours and scan.
4. Prophetic characterization of viral maturation polypeptides and systems

The methods disclosed herein are useful for, among other things, identifying host proteins involved in virus release. To this end we compare membrane protein profiles of cells infected with viruses expressing wild type (wt) p6Gag (p6) with the parallel profiles from cells infected with viruses harboring a mutant form of p6. Wt p6 is expected to attract the release machinery and it is expected that the mutant virus fails to do so. Furthermore, we

also investigate the mechanism of the exceedingly high efficiency release mechanism of the Ebola virus by comparing the protein profiles of membranes from cells infected with mutant HIV-1 expressing the Ebola release determinant with membranes of cells infected with wt virus.

Host protein profiling

Membranes from uninfected, wt virus and mutant virus-infected cells are prepared according to protocols modified to enable virus inactivation prior to sample handling and separation. The membranes samples are separated by 2D gel electrophoresis (2DGE). The 2D maps are analyzed and proteins specific to wt virus are subjected to mass spectrometry analysis.

Virion protein profiling

Host proteins involved in virus release are expected to be trapped in virions after their release. This expectation is based, in part, on two observations. The first is the presence of ubiquitin in virions at concentrations higher than the ubiquitin cellular concentration. The second is the finding that EIAV gag protein associates with AP50 and that AP-2 is also found in EIAV particles. We have found mono-ubiquitinated p9 of EIAV in the virions.

It is therefore possible that additional host proteins are included in virions. Analysis of host proteins included in virus particles facilitates identification of host proteins involved in virus budding and release.

To identify host proteins included in virus particles, we harvest the virions from the supernatant of virus-infected cells. The virions are then lysed and the proteins are subsequently analyzed by 2DGE. Specific antibodies are used to identify the viral proteins. The unidentified host proteins are subjected to MS analysis. Antibodies to the host proteins present in the virus particles are used to detect them in membranes of virus infected cells. It is believed that host proteins present in virions and localization in sites of virus budding will be involved in virus maturation.

Identification of ubiquitinated proteins associated with virus release

Cells transfected with hemagglutinin (HA)-tagged ubiquitin are infected with the relevant virus. To isolate the ubiquitinated proteins, detergent lysates are prepared and the

detergent extract is subjected to immunoprecipitation with anti-HA antibody. The immunoprecipitates are subjected to separation by either SDS-PAGE or 2DGE (depending on the complexity of the proteome). Using this approach we compare ubiquitinated proteins from wt and mutant virus-infected cells. Those proteins that are ubiquitinated in wt but not in mutant-infected cells are identified and characterized by mass spectrometry analysis.

Identification of ubiquitin-protein ligases

The rate-limiting component for virus release is expected to be a ubiquitin-protein ligase. The Ebola recruits the ligase to the sites of budding with exceedingly higher efficiency than any other retrovirus.

1. We generate a recombinant HIV1-p6 with the Ebola tandem L motif. To confirm the Ebola potency we compare virus-like particle release into the medium of cells expressing the two p6 forms. Virus-like particles are harvested from the medium of p6-expressing cells. We quantitatively detect Gag with specific antibodies by immunoblot analysis. It is expected that Gag signal will be much higher in the supernatant of p6_{Ebola} expressing cells.
2. We prepare membraned protein-enriched fraction. To this end we prepare protein fractions from unwashed or mildly washed membrane so as to minimally disturb possible ligase-protein interactions. Protein profiles of plasma membrane proteins from un-infected, p6_{HIV-1} and P6_{Ebola}-transfected cells are compared. Proteins that are expressed at higher levels in the P6 Ebola membranes are analyzed and identified by mass spectrometry. We specifically search for proteins with the characteristics of an AVMSp.
3. Alternatively, we utilize anti-p6 antibodies to precipitate proteins that are associated with p6. A 2D profile of p6-associated proteins is performed and results are analyzed via a similar rationale as described in part 2 above.

INCORPORATION BY REFERENCE

All publications and patents mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

EQUIVALENTS

While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of this specification and the claims below. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

We Claim:

1. An isolated protein complex comprising a RING-SH₃ polypeptide in combination with at least one polypeptide selected from the group consisting of: a Gag protein, a Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin, RING-SH₃, and a clathrin.
2. The isolated protein complex of claim 1, wherein said Gag protein is an HIV gag protein.
3. The isolated protein complex of claim 1, wherein said Gag protein comprises the Gag late domain.
4. The isolated protein complex of claim 3, wherein said Gag late domain is PTAP.
5. The isolated protein complex of claim 3, wherein said Gag late domain is PxxY.
6. The isolated protein complex of claim 3, wherein said Gag late domain is PxxL.
7. The isolated protein complex of claim 3, wherein said Gag late domain is PPxY.
8. The isolated protein complex of claim 3, wherein said Gag late domain is YxxL.
9. The isolated protein complex of claim 3, wherein said Gag late domain is PxxP.
10. A host cell comprising a first nucleic acid and a second nucleic acid, wherein the first nucleic acid comprises a recombinant RING-SH₃ nucleic acid, and wherein the second nucleic acid comprises a recombinant nucleic acid encoding a Gag protein.
11. The host cell of claim 10, wherein said Gag protein is an HIV gag protein.
12. The host cell of claim 11, wherein said Gag protein comprises the Gag late domain.
13. The host cell of claim 12, wherein said Gag late domain is PTAP.
14. The host cell of claim 12, wherein said Gag late domain is PxxY.

15. The host cell of claim 12, wherein said Gag late domain is PxxL.
16. The host cell of claim 12, wherein said Gag late domain is PPxY.
17. The host cell of claim 12, wherein said Gag late domain is YxxL.
18. The host cell of claim 12, wherein said Gag late domain is PxxP.
19. A method for identifying modulators of protein complexes, comprising:
 - (i) forming a reaction mixture comprising
 - (a) a RING-SH₃; and
 - (b) a second polypeptide selected from the group consisting of: RING-SH₃, a gag protein, a Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin, and a clathrin;
 - (ii) contacting said reaction mixture with a test agent, and
 - (iii) determining the effect of said test agent for one or more activities selected from the group comprising (a) a change in the level of the protein complex, (b) a change in the enzymatic activity of the complex, or (c) where the reaction mixture is a whole cell, a change in the plasma membrane localization of the complex or a component thereof.
20. A method for identifying a test compound which inhibits or potentiates complex formation, comprising:
 - (i) forming a reaction mixture comprising
 - (a) a RING-SH₃; and
 - (b) a second polypeptide selected from the group consisting of: RING-SH₃, a gag protein, a Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a GTPase, an E2 enzyme, tsg101, a cullin, and a clathrin;
 - (ii) contacting said reaction mixture with a test agent, and
 - (iii) detecting binding of said RING-SH₃ to said second polypeptide;wherein a change in the binding of said RING-SH₃ to said second polypeptide in the presence of the test compound, relative to binding in the absence of the test compound, indicates that said test compound potentiates or inhibits complex formation between said RING-SH₃ and said second polypeptide.

21. A method for inhibiting infection in a subject in need thereof, comprising administering an effective amount of an agent that inhibits the binding of a RING-SH₃ polypeptide to an gag protein.
22. The method of claim 21, wherein said agent is selected from the group comprising a small molecule, a antibody, and a peptide.
23. The method of claim 22, wherein the Gag protein is an HIV Gag protein.
24. The method of claim 22, wherein the Gag polypeptide is HIV p24.
25. An isolated antibody, or fragment thereof, specifically immunoreactive with an epitope of a RING-SH₃ polypeptide, which disrupts the interaction between said RING-SH₃ and a RING-SH₃-associating polypeptide (RING-SH₃-AP).
26. The antibody of claim 25, wherein said antibody is a monoclonal antibody.
27. The antibody of claim 25, wherein said antibody is a Fab fragment.
28. The antibody of claim 25, wherein said antibody is labeled with a detectable label.
29. The antibody of claim 25, wherein said RING-SH₃-AP is a gag polypeptide.
30. The antibody of claim 25, wherein said RING-SH₃-AP is an HIV gag polypeptide.
31. A kit for detecting a RING-SH₃ polypeptide protein comprising (i) isolated anti- RING-SH₃ antibodies, or fragment thereof, specifically immunoreactive with an epitope of an RING-SH₃, which epitope interacts with an RING-SH₃-AP, and (ii) a detectable label for detecting said anti- RING-SH₃ antibody in immunoclonplexes with said RING-SH₃ polypeptide.
32. A host cell comprising a first nucleic acid and a second nucleic acid, wherein the first nucleic acid comprises a recombinant RING-SH₃ nucleic acid, and wherein the second nucleic acid comprises a recombinant nucleic acid encoding a HIV Gag protein.

33. A method for inhibiting infection in a subject in need thereof, comprising administering an effective amount of an agent that inhibits the binding of a RING-SH₃ polypeptide to an HIV gag protein.
34. An anti-viral therapeutic composition comprising a double stranded oligoribonucleotide molecule that inhibits expression of a nucleic acid molecule encoding a RING-SH₃ polypeptide

Figure 1: Human PRT3 Amino Acid Sequence (SEQ ID NO:1)

MDESALLDLECPVCLERLDASAKVLPCHTFCKRCLLGIVGSRNELRCPECRTLVGSGVEELPSNILLV
RLLDGIKQRPWKPGPGGGSGTNCNALSQSSTVANCSSKDLQSSQGGQQPRVQSWSPVVRGIPQLPCAK
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFFPTNFVQIIKPLPQPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIRRVNENWAEGLADKIGIFPISYVEFNAAKQLIEWDKPPVPGVDAGECS
SAAQSSSTAPKHSDDTKKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHISTTGLIVTPPPSSPVTTGPFSTFPDVPYQAALGTLNPPLPPPPLLAATVLASTPPGATAA
AAAAGMGPRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHT
SKIGVFPNGYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQGNVAGSPSVVPAVV
SAAHIQTSPQAKVLLHMTGQMTVNQARNAVRTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLAS
PQPAPLMPGSATHTAASISRASAPLACAAAAPLTSPSITSASLEAEPSGRIVTVLPGLPTSPDSASSAC
GNSSATKPKDKSKKEKKGLLKLKLSGASTKRKPRVSPASPTLEVELGSAELPLQGAVGPELPPGGGHGRA
GSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAPPPRQACSSLGPVLNESRPVVCERHRVVVSY
PPQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLFPGSFVENI

Figure 2A: N terminus protein fragment of hPOSH (public gi:10432612; SEQ ID NO:2)

MDESALLDLLECPVCLERLDASAKVLPCQHTFCKRCLLGI VGSRNELRCPECRTL VSGSGVEELPSNILLV
RLLDGIKQRPWKPGPGGGSGTNCTNALRSQSSTVANCSSKDLQSSQGGQQPRVQSWSPVVRGIPQLPCA
ALYNYEGKEPGDLKFSKGDIIILRRQVDENWYHGEVNGIHGFFPTNFVQIIKPLPQPPPQCKALYDFEVK
DKEADKDCLPFAKDDVLTVIRRV DENWAEGLADKIGIFPISYVEFN SAAKQLIEWDKPPVPGVDAGECS
SAAQSSSTAPKHSDTKKNTKKRHSFTSLTMANKSSQASQNRHSMEISPPVLISSSNPTAAARISELSGLS
CSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFP SDVPYQAALGTLNPPPLPPPPLLAATV LASTPPGATAA
AAAAAGMGRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPRKEDELELRKGEMFLVFERCQDGWFKGTSMHT
SKI GVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVTMVSPSTAGGPAQKLQGNVAGSPSVVPAAVV
SAAHIQTSPQAKVLLHMTGQMTVNQARNVRTVAAHNQERPTAAVTPIQVQNAAGLSPASVGLSHHSLAS
PQPAPLMPGSATHTA AISISRASAPLACAAAAPLTSPSITSASLEAEPSGRIVTVLPGLPTSPDSASSAC
GNSSATKPKDKDS

Figure 2B: C terminus protein fragment of hPOSH (public gi:7959249) (SEQ ID NO:39)

ISYVEFN SAAKQLIEWDKPPVPGVDAGECSSAAQSSSTAPKHSDTKKNTKKRHSFTSLTMANKSSQASQ
NRHSMEISPPVLISSSNPTAAARISELSGLSCSAPSQVHISTTGLIVTPPPSSPVTTGPSFTFP SDVPYQA
ALGTLNPPPLPPPPLLAATV LASTPPGATAAAAAAGMGRPMAGSTDQIAHLRPQTRPSVYVAIYPYTPR
KEDELELRKGEMFLVFERCQDGWFKGTSMHTSKI GVFPGNYVAPVTRAVTNASQAKVPMSTAGQTSRGVT
MVSPSTAGGPAQKLQGNVAGSPSVVPAAVVSA AHIQTSPQAKVLLHMTGQMTVNQARNVRTVAAHNQER
PTAAVTPIQVQNAAGLSPASVGLSHHSLASPQPAPLMPGSATHTA AISISRASAPLACAAAAPLTSPSIT
SASLEAEPSGRIVTVLPGLPTSPDSASSACGNSSATKPKDKDSKKEKKGLLKL LSGASTKRKPRVSPPASP
TLEVELGSAELPLQGA VGPELPPGGGHGRAGSCPVDGDGPVTTAVAGAALAQDAFHRKASSLDSAVPIAP
PPRQACSSLGPVLNESRPVVCERHRVVVSYP PQSEAELELKEGDIVFVHKKREDGWFKGTLQRNGKTGLF
PGSFVENI

Figure 3

Name: Similar to MUS
Musculus Plenty of
SH3S (PRT3)

Gen Bank Accession No. BAB13822.1

SEQ ID No. 3

SEQUENCE 712 AA

```
MDESALLDLL ECPVCLERLD ASAKVLPCQH TFCKRCLLGI VGSRNELRCP ECRTLVGSGV
EELPSNILLV RLLDGIKQRP WKPGPGGGSG TNCTNALRSQ SSTVANCSSK DLQSSQGGQQ
PRVQSWSPPV RGIPQLPCAK ALYNYEGKEP GDLKFSKGDI IILRRQVDEN WYHGEVNGIH
GFFPTNFVQI IKPLPQPPPQ CKALYDFEVK DKEADKDCLP FAKDDVLTVI RRVNENWAEQ
MLADKIGIFP ISYVEFNAA KQLIEWDKPP VPGVDAGECS SAAQSSSTAP KHSDDTKKNTK
KRHSFTSLTM ANKSSQASQN RHMELSPPV LISSSNPTAA ARIEELSGLS CSAPSQVHIS
TTGLIVTPPP SSPVTTGPSF TFPDVPYQA ALGTLNPPLP PPPLLAATVL ASTPPGATAA
AAAAGMGPRP MAGSTDQIAH LRPQTRPSVY VAIYPYTPRK EDELELRKGE MFLVFERCQD
GWFKGTSMHT SKIGVFPQNY VAPVTRAVTN ASQAKVPMST AGQTSRGVTM VSPSTAGGPA
QKLQNGVAG SPSVVPAAVV SAAHIQTSPQ AKVLLHMTGQ MTVNQARNAV RTVAAHNQER
PTAAVTPIQV QNAAGLSPAS VGLSHHSLAS PQPAPLMPGS ATHTAAISIS RASAPLACAA
AAPLTSPSIT SASLEAEPST RIVTVLPGLP TSPDSASSAC GNSSATKPKD DS
```

Figure 3

Figure 4

Name: SH3-BINDING
PROTEIN CBL-B

Gen Bank Accession No. AAB09291.1 No: 4

SEQUENCE 982 AA

```
MANSMNGRNP GGRGGNPRKG RILGIIDAIQ DAVGPPKQAA ADRRTVEKTW KLMDKVVRLC
QNPQLQLKNS PPYILDILPD TYQHLRLILS KYDDNQKLAQ LSENEYFKIY IDSLMKKSKR
AIRLFKEGKE RMYEEQSQDR RNLTKLSLIF SHMLAEIKAI FPNGQFQGDN FRITKADAAE
FWRKFFGDKT IVPWKVFRQC LHEVHQISSS LEAMALKSTI DLTCNDYISV FEEDIFTRLF
QPWGSILRNW NFLAVTHPGY MAFLTYDEVK ARLQKYSTKP GSYIFRLSCT RLQQAIGYV
TGDGNILQTI PHNKPLFQAL IDGSREGFYI YPDGRSYNPD LTGLCEPTPH DHIKVTQEYQ
ELYCEMGSTF QLCKICAEND KDVKIEPCGH LMCTSCLTAW QESDGGGCPF CRCEIKGTEP
IIVDPFDPD EGSRCCSIID PFGMPMLDLD DDDREESLM MNRLANVRKC TDRQNSPVTs
PGSSPLAQRK KPQPDPLQIP HLSLPPVPPR LDLIQKGIVR SPCGSPTGSP KSSPCMVRKQ
DKPLPAPPPP LRDP PPPPPE RPPPIPPDNR LSRHHHVES VPSRDPPMPL EAWCPRDVFG
TNQLVGCRLI GEGSPKPGIT ASSNVNGRHS RVGSDPVLMR KHRHDLPLE GAKVFSNGHL
GSEEDVPPR LSPPPPVTTL LPSIKCTGPL ANSLSEKTRD PVEEDDDEYK IPSSHVSLN
SQPSHCHNVK PPVRSCDNH CMLNGTHGPS SEKKSNIPDL SIYLGKDVFD SASDPVPLPP
ARPPTRDNPK HGSSLNRTPS DYDLLIPPLG EDAFDALPPS LPPPPPPARH SLIEHSKPPG
SSSRPSSGQD LFLLPDPFV DLASGQVPLP PARRLPGENV KTNRTSQDYD QLPSCSDGSQ
APARPPKPRP RRTAPEIHRH KPHGPAALE NVDAKIAKLM GEGYAFEEVK RALEIAQNNV
EVARSILREF AFPPPVSPRL NL
```

Figure 4

Figure 5

Name: SH3-BINDING
PROTEIN CBL-C

Gen Bank Accession No. BAA86298.1

SEQ ID No. 5

SEQUENCE 474 AA

```
MALAVAPWGR QWEEARALGR AVRMLQRLEE QCVDPRLSVS PPSLRDLLPR TAQLLREVAH
SRRAAGGGGP GPGGSGDFL LIYLANLEAK SRQVAALLPP RGRRSANDEL FRAGSRLRRQ
LAKLAIIFSH MHAELHALFP GGKYCGHMYQ LTKAPAHTFW RESCGARCVL PWAEFESLLG
TCHPVEPGCT ALALRTTIDL TCSGHVSIFE FDVFTRLFQP WPTLLKNWQL LAVTHPGYMA
FLTYDEVQER LQACRDKPGS YIFRPCTRL GQWAIGYVSS DGSILQTIPA NKPLSQVLLE
GQKDGFYLYP DGKTHNPDLT ELGQAEPPQR IHVSEEQLQL YWAMDSTFEL CKICAESNKD
VKIEPCGHLL CSCCLAAWQH SDSQTCPFCR CEIKGWEAVS IYQFYGQATA EDSGNSSDQE
GRELELGQVP LSAPPLPPRP DLPPRKPRNA QPKVRLKGN SPPAALGPQD PAPA
```

Figure 5

Figure 6

Name: Multiple Membrane
Spanning Receptor TRC8

Gen Bank Accession No. AAC39930.1

SEQ ID No. 6

SEQUENCE 664 AA

```
MAAVGPPQQQ VRMAHQIWA ALEVALRVPC LYIIDAIFNS YPDSSQSRFC IVLQIFLRLF
GVFASSIVLI LSQRSLFKFY TYSSAFLLAA TSVLVNYYAS LHIDFYGAYN TSAFGIELLP
RKGPSLWMAL IVLQLTFGIG YVTLLQIHSI YSQLIILDLL VPVIGLITEL PLHIRETLFF
TSSLILTLNT VFVLAVKLKW FYYSTRYVYL LVRHMYRIYG LQLLMEDTWK RIRFPDILRV
FWLTRVTAQA TVLMYILRMA NETDSFFISW DDFWDLICNL IISGCDSTLT VLGMSAVISS
VAHYLGILGIL AFIGSTEEDD RRLGFVAPVL FFILALQTGL SGLRPEERLI RLSRNMCLLL
TAVLHFIHGM TDPVLMSLSA SHVSSFRRHF PVLFSACL F ILPVLLSYVL WHHYALNTWL
FAVTAFCVEL CLKVIVSLTV YTLFMIDGY NVLWEKLDDY VYYVRSTGSI IEFIFGVVMF
GNGAYTMMFE SGSKIRAFMM CLHAYFNIYL QAKNGWKTFM NRRTAVKKIN SLPEIKGSRL
QEINDVCAIC YHEFTTSARI TPCNHYFHAL CLRKWLYIQD TCPMCHQKVY IEDDIKDNSN
VSNNGFIPP NETPEEAVRE AAAESDRELN EDDSTDCDD VQRERNVIO HTGAAAEFFN
DDTD
```

Figure 6

Figure 7

Name: Autocrine Motility
Factor Receptor

Gen Bank Accession No. AAD56722.1

SEQ ID No. 7

SEQUENCE 643 AA

```
MPLLFLERFP WPSLRITYTGL SGLALLGTII SAYRALSQPE AGPGEFDQLT ASLQPEPPAP
ARPSAGGPRA RDVAQYLLSD SLFVWVLVNT ACCVLMLVAK LIQCIVFGPL RVSERQHLKD
KFWNFIFYKF IFIFGVINVQ TVEEVVMWCL WFAGLVFLHL MVQLCKDRFE YLSFSPTTPM
SSHGRVLSLL VAMLLSCCGL AAVCSITGYT HGMHTLAFMA AESLLVTVRT AHVILRYVIH
LWDLNHEGTW EGKGTYYVYT DFMELTLLS LDLMHIIHML LFGNIWLSMA SLVIFMQLRY
LFHEVQRRIR RHKNYLRVVG NMEARFAVAT PEELAVNND CAICWDSMQA ARKLPCGHLF
HNSCLRSWLE QDTSCPTCRM SLNIADNNRV REEHQGENLD ENLVPVAAAE GRPRLNQHNH
FFHFDGSRIA SWLPFSFSEV MHTNINLGIT QASNSQLNAM AHQIQEMFPQ VPHYHLVLQDL
QLTRSVEITT DNILEGRIQV PFPTQRSDSI RPALNSPVER PSSDQEEGET SAQTERVPLD
LSPRLEETLD FGEVEVEPSE VEDFEARGSR FSKSADERQR MLVQRKDELL QQARKRFLNK
SSEDDAASES FLPLEGASSD PVTLRRRMLA AAAERRLQKQ QTS
```

Figure 7

Figure 8

Name: Membrane-Associated Nucleic Acid Binding Protein (Fragment) Gen Bank Accession No. AAG00432.1

SEQ ID No. 8

SEQUENCE 1048 AA

```
MPVQAAQWTE FLSCPICYNE FDENVHKPIS LGCSHTVCKT CLNKLHRKAC PFDQTAINTD
IDVLPVNFAL LQLVGAQVPD HQSIKLSNLG ENKHYEVAKK CVEDLALYLK PLSGGKGVAS
LNQSALSRPM QRKLVTLVNC QLVEEEGRVR AMRAARSLGE RTVTELILQH QNPQQLSANL
WAAVRARGCQ FLGPAMQEEA LKLVLLELED GSALSRKVLV LFVVQRLEPR FPQASKTSIG
HVVQLLYRAS CFKVTKRDED SSLMQLKEEF RSYEALRREH DAQIVHIAME AGLRISPEQW
SSLLYGDLAH KSHMQSIIDK LQSPESFAKS VQELTIVLQR TGD PANLNRL RPHLELLANI
DPNPDAVSPT WEQLENAMVA VKTVVHGLVD FIQNYSRKGH ETPQQPQNSK YKTSMCRDLR
QQGGCPRGTN CTFAHSQEEL EKYRLRNKKI NATVRTFPLL NKVGVNNTVT TTAGNVISVI
GSTETTGTKIV PSTNGISNAE NSVSQLISRS TDSTLRALET VKKVGVKVGAN GQNAAGPSAD
SVTENKIGSP PKTPVSNVAA TSAGPSNVGT ELNSVPQKSS PFLTRVPVYP PHSENIQYFQ
DPRTQIPFEV PQYPQTGYYP PPPTVPAGVA PCVPRFVRSN NVPESSLPPA SMPYADHYST
FSPRDRMNSS PYQPPPPQPY GPVPPVPSGM YAPVYDSRRI WRPPMYQRDD IIRSNLPPM
DVMHSSVYQT SLRERYNSLD GYYSVACQPP SEPRTTVPLP REPCGHLKTS CEEQIRRKPD
QWAQYHTQKA PLVSSTLPVA TQSPTPPSPL FSVDFRADFS ESVSGTKFEE DHLSHYSPWS
CGTIGSCINA IDSEPKDVIA NSNAVLMDLD SGDVKRRVHL FETQRRRTKEE DPIIPFSDGP
IISKWGAISR SSRTGYHTTD PVQATASQGS ATKPISVSDY VPYVNAVDSR WSSYGNEATS
SAHYVERDRF IVTDLSGHRK HSSTGDLLSL ELQQAKSNSL LLQREANALA MQQKWNLSLDE
GRHLTLNLLS KEIELRNGEL QSDYTEDA
```

Figure 8

Figure 9

Name: NADH- Ubiquinone Gen Bank Accession No. CAA57489.1 SEQ ID No. 9
Oxidoreductase MWFE Subunit

SEQUENCE 70 AA;

MWFEILPGLS VMGVCLLIPG LATAYIHRFT NGGKEKRVAH FGYHWSLMER DRRISGVDRY
YVSKGLENID

Figure 9

Figure 10

Name: Peroxisome Assembly
Factor-1 (PAF-1)

Gen Bank Accession No. AAC12785.1

SEQ ID No. 10

SEQUENCE 305 AA

MASRKENAKS ANRVLRLISQL DALELNKALE QLVWSQFTQC FHGFKPGLLA RFEPEVKACL
WVFLWRFTIY SKNATVGQSV LNIKYKNDFS PNLRYQPPSK NQKIWYAVCT IGGRWLEERC
YDLFRNHHLA SFGKVKQCVN FVIGLLKLGG LINFLIFLQR GKFATLTERL LGIHSVFCKP
QNIREVGFY MNRELLWHGF AEFLIFLLPL INVQKLKAKL SSWCIPLTGA PNSDNTLATS
GKECALCGEW PTMPHTIGCE HIFCYFCAKS SFLFDVYFTC PKCGTEVHSL QPLKSGIEMS
EVNAL

Figure 10

Figure 11

Name: Peroxisome Assembly
Protein 10

Gen Bank Accession No. AAC18133.1

SEQ ID No. 11

SEQUENCE 326 AA

```
MAPAAASPPE VIRAAQKDEY YRGGLRSAAG GALHSLAGAR KWLEWRKEVE LLSDVAYFGL
TTLAGYQTLG EEYVSIIQVD PSRIHVPSSL RRGVLVTLHA VLPYLLDKAL LPLEQELQAD
PDSGRPLQGS LGPGGRGCSG ARRWMRHHTA TLTEQORRAL LRAVFVLRQG LACLQRLHVA
WFYIHGVFYH LAKRLTGITY LRVRSLPGED LRARVSYRL GVISLLHLVL SMGLQLYGFR
QRQRARKEWR LHRGLSHRRA SLEERAVSRN PLCTLCLEER RHPTATPCGH LFCWECITAW
CSSKAECPLC REKFPPQKLI YLRHYR
```

Figure 11

Figure 12

Name: Peroxisome Assembly
Protein 12

Gen Bank Accession No. AAC68812.1

SEQ ID No. 12

SEQUENCE 359 AA

```
MAEHGAHFTA ASVADDQPSI FEVVAQDSLM TAVRPALQHV VKVLAESNPT HYGFLWRWFD
EIFTLDDLIL QQHYLSRTSA SFSENFYGLK RIVMGDTHKS QRLASAGLPK QQLWKSIMFL
VLLPYLKVKL EKLVSLLREE DEYSTHPPSS RWKRFYRAFL AAYPFVNMAW EGWFLVQQLR
YILGKAQHHS PLLRLAGVQL GRLTVQDIQA LEHKPAKASM MQQPARSVSE KINSALKKAV
GGVALSLSTG LSVGVFLLQF LDWWYSSENQ ETIKSLTALP TPPPPVHLDY NSDSPLLPKM
KTVCPICRKT RVNDTVLATS GYVFCYRCVF HYVRSHQACP ITGYPTFVQH LIKLYSPEN
```

Figure 12

Figure 13

Name: Acetylcholine Receptor- Gen Bank Accession Nos. CAA83954.1 SEQ ID No: 13
Associated 43 KDA Protein

SEQUENCE 412 AA

```
MGQDQTKQQI EKGLQLYQSN QTEKALQVWT KVLEKSSDLM GRFRVLGCLV TAHSEMGRYK
EMLKFAVVQI DTARELEDAD FLLESYLNLA RSNEKLCEFH KTISYCKTCL GVPGTRAGAQ
LGGQVSLSMG NAFLGLSVFQ KALESFEKAL RYAHNNDDTM LECRVCCSLG SFYAQVKDYE
KALFFPCKAA ELVNNYGKGW SLKYQAMSQY HMAVAYRLLG RLGSAMECCE ESMKIALQHG
DRPLQALCLL CFADIHRSRG DLETAFPRYD SAMSIMTEIG NRLGQVQALL GVAKCWVARK
ALDKALDAIE RAQDLAEEVG NKLSQLKLHC LSESIYRSKG LQRELRAHVV RFHECVEETE
LYCALCGESI GEKNSRLQAL PCSHIFHLRC LQNNGTRSCP NCRRSSMKPG FV
```

Figure 13

Figure 14

Name: Peroxisomal Membrane
Protein 3

Gen Bank Accession No. AAH00661.1

SEQ ID No 14

SEQUENCE 305 AA

MASRKENAKS ANRVLRI SQL DALELNKALE QLVWSQFTQC FHGFKPGLLA RFEPEVKACL
WVFLWRFTIY SKNATVGQSV LNIKYKNDFS PNLRYQPPSK NQKIWYAVCT IGGRWLEERC
YDLFRNHHLA SFGKVKQCVN FVIGLLKLGG LINFLIFLQR GKFATLTERL LGIHSVFCKP
QNIREVGFY MNRELLWHGF AEFLIFLLPL INVQKLKAKL SSWCIPLTGA PNSDNTLATS
GKECALCGER PTMPHTIGCE HIFCYFCAKS SFLFDVYFTC PKCGTEVHSL QPLKSGIEMS
EVNPL

Fig. 14

Figure 15

Name: HsRmalGen Bank Accession No. BAB39359.1SEQ ID No 15

SEQUENCE 180 AA

MAAAEEEDGG PEGPNRERGG AGATFECNIC LETAREAVVS VCGHLYCWPC LHQWLETRPE
RQECFVCKAG ISREKVVPLY GRGSQKPQDP RLKTPPRPQG QRPAPESRGG FQPFQDTGGF
HFSFGVGAFP FGFFTTFVNA HEPFRRGTGV DLGQGHPASS WQDSLFLFLA IFFFFWLLSI

Figure 15

Figure 16

Name: HtripGen Bank Accession No. AAB52993.1SEQ ID No 16

SEQUENCE 469 AA

```
MPIRALCTIC SDFFDHSRDV AAIHCGHTFH LQCLIQSFET APSRTCPCQR IQVGKRTIIN
KLFFDLAQEE ENVLDREFLK NELDNVRAQL SQKDKEKRDS QVIIDTLRDT LEERNATVVS
LQQALGKAEM LCSTLKKQMK YLEQQQDETK QAQEEAGRRLR SKMKTMEQIE LLLQSQLPEV
EEMIRDMGVG QSAVEQLAVY CVSLKKEYEN LKEARKASGE VADKLRKDLF SSRSKLQTVY
SELDQAKLEL KSAQKDLQSA DKEIMSLKKK LTMLQETLNL PPVASETVDR LVLESPAPVE
VNLKLRRPSF RDDIDLNATF DVDTPPARPS SSQHGYEKL CLEKSHSPIQ DVPKKICKGP
RKESQLSLGG QSCAGEPDEE LVGAFFIFVR NAILGQKQPK RPRSESSCSK DVVRTGFDGL
GGRTKFIQPT DTVMIRPLPV KPCTKVQQRV RVKTVPSLFQ AKLDTFLWS
```

Figure 16

Figure 17

Name: Cystein Rich Domain
Associated To Ring and Traf
Protein

Gen Bank Accession No. CAA56491.1

SEQ ID No 17

SEQUENCE 470 AA

```
MPGFDYKFLE KPKRLLCPL CGKPMREPVO VSTCGHRFCD TCLQEFLSEG VFKCPEDQLP
LDYAKIYPDP ELEVQVLGLP IRCIHSEEGC RWSGPLRHLQ GHLNTCSFNV IPCPNRCPMK
LSRRDLPAHL QHDCPKRRLK CEF CGCDFSG EAYESHEGMC PQESVYCENK CGARMMRGLL
AQHATSECPK RTQPCTYCTK EFVFDTIQSH QYQCPRLPVA CPNQCGVGT V AREDLPGHLK
DSCNTALVLC PFKDSGCKHR CPKLAMARHV EESVKPHLAM MCALVSRQRQ ELQELRRELE
ELSVGSDGVL IWKIGSYGRR LQEAKAKPNL ECFSPAFYTH KYGYKLQVSA FLNGNGSGEG
THLSLYIRVL PGAFDNLLEW PFARRVT FSL LDQSDPGLAK PQHVTET FHP DPNWKNFQKP
GTWRGSLDES SLGFGYPKFI SHQDIRK RNY VRDDAVFIRA AVELPRKILS
```

Figure 17

Figure 18

Name: PML-RARGen Bank Accession No. AAA60126.1SEQ ID No 18

SEQUENCE 797 AA

```
MEPAPARSPR PQQDPARPQE PTMPPPETPS EGRQSPSPSPS PTERAPASEE EFQFLRCQQC
QAEAKCPKLL PCLHTLCSGC LEASGMQCPI CQAPWPLGAD TPALDNVFFE SLQRRLSVYR
QIVDAQAVCT RCKESADFWC FECEQLLCAK CFEAHQWFLK HEARPLAELR NQSVREFLDG
TRKTNNIFCS NPNHRTPTLT SIYCRGCSKP LCCSCALLDS SHSELKCDIS AEIQQRQEEL
DAMTQALQEQ DSAFGAVHAQ MHAAVGQLGR ARAETEELIR ERVRQVVAHV RAQERELLEA
VDARYQRDYE EMASRLGRLD AVLQRI RTGS ALVQRMKCYA SDQEVLDMHG FLRQALCRLR
QEEPQSLQAA VRTDGFDEFK VRLQDLSSCI TQGKAIETQS SSSEEIVPSP PSPPPLPRIY
KPCFVCQDKS SGYHYGVSAC EGCKGFFRRS IQKNMVYTCH RDKNCIINKV TRNRCQYCRL
QKCFEVGMSK ESVRNDRNKK KKEVPKPECS ESYTLTPEVG ELIEKVRKAH QETFPALCQL
GKYTTNNSSE QRVSLDIDLW DKFSELSTKC IIKTVEFAKQ LPGFTTLTIA DQITLLKAAC
LDILILRICT RYTPQD TMT FSDGLTLNRT QMHNAGFGPL TDLVFAFANQ LLPLEMDDAE
TGLLSAICLI CGDRQDLEQP DRVDMLEQPL LEALKVYVRK RRPSRPHMFP KMLMKITDLR
SISAKGAERV ITLKMEIPGS MPPLIQEMLE NSEGLDTLSG QPGGGGRDGG GLAPPPGSCS
PSLSPSSNRS SPATHSP
```

Figure 18

Figure 19

Name: TRAF5Gen Bank Accession No. BAA25262.1SEQ ID No 19

SEQUENCE 557 AA

MAYSEEHKGM PCGFIRQNSG NSISLDFEPS IEYQFVERLE ERYKCAFCHS VLHNPHQTGC
GHRFCQHCIL SLRELNTVPI CPVDKEVIKS QEVFKDNCCK REVLNLYVYC SNAPGCNAKV
ILGRYQDHLQ QCLFQPVQCS NEKCREPVLK KDLKEHLSAS CQFRKEKCLY CKKDVVVINL
QNHEENLCPE YPVFCPNNCA KIILKTEVDE HLAVCPAEQ DCPFKHYGCA VTDKRRNLQQ
HEHSALREHM RLVLEKNVQL EEQISDLHKS LEQKESKIQQ LAETIKKLEK EFKQFAQLFG
KNGSFLPNIQ VFASHIDKSA WLEAQVHQLL QMVNQQQNKF DLRPLMEAVD TVKQKITLLE
NNDQRLAVLE EETNKHDTHI NIHKAQLSKN EERFKLLEGT CYNGKLIWKV TDYKMKKREA
VDGHTVSIFS QSFYTSRCGY RLCARAYLNG DSGRGSLS LYFVVMRGEF DSSLQWPPFRQ
RVTLMLLDQS GKKNIMETFK PDPNSSSFKR PDGEMNIASG CPRFVAHSV L ENAKNAYIKD
DTLFLKVAVD LTDLEDL

Figure 19

Figure 20

Name: Tumor Necrosis Factor
Receptor-Associated
Factor 4A

Gen Bank Accession No. AAC32376.1

SEQ ID No 20

SEQUENCE 198 AA

MPGFDYKFLE KPKRRLLCPL CGKPMREPVO VSTCGHRFCD TCLQEFLLSEG VFKCPEDQLP
LDYAKIYPDP ELEVQVLGLP IRCIHSEEGC RWSGPLRHLQ GHLNTCSFNV IPCPNRCPMK
LSRRDLPAHL QHDCPKRRLK CEF CGCDFSG EAYESHESL GFGYPKFISH QDIRKRNRYVR
DDAVFIRA AV ELPRKILS

Figure 20

Figure 21

Name: HFB30Gen Bank Accession No. BAA78677.1SEQ ID No 21

SEQUENCE 474 AA

```
MSEDREAQE DELLALASIY DGDEFKRAES VQGGETRIYL DLPQNFKIFV SGNSNECLQN
SGFEYTTICFL PPLVLNFELP PDYPSSSPPS FTLSGKWLSP TQLSALCKHL DNLWEEHRGS
VVLFAWMQFL KEETLAYLNI VSPFELKIGS QKKVQRRTAQ ASPNTELDGF GAAGSDVDQE
EIVDERAVQD VESLSNLIQE ILDFDQAQOI KCFNSKLFLC SICFCEKLGs ECMYFLECRH
VYCKACLKDY FEIQIRDGQV QCLNCPEPKC PSVATPGQVK ELVEAELFAR YDRLLLQSSL
DLMADVVCYP RPCCQLPVMQ EPGCTMGICS SCNFAFCTLG RLTYHGVSPC KVTAEKLMDL
RNEYLQADEA NKRLLDQRYG KRVIQKALEE MESKEWLEKN SKSCPCCGTP IEKLDGCNKM
TCTGCMQYFC WICMGSLSRA NPYKHFNDPG SPCFNRLFYA VDVDDDIWED EVED
```

Figure 21

Figure 22

Name: BIR2Gen Bank Accession No. ACC41943.1SEQ ID No 22

SEQUENCE 604 AA

```
MNIVENSIFL SNLMKSANTF ELKYDLSCEL YRMSTYSTFP AGVPVSERSL ARAGFYITGV
NDKVKCFCCG LMLDNWKRGD SPTEKHKKLY PSCRFBVQSLN SVNNLEATSQ PTFPSSVTNS
THSLLPGTEN SGYFRGSYSN SPSNPVNSRA NQDFSALMRS SYHCAMNNEN ARLLTFQTWP
LTFLSPTDLA KAGFYIIGPG DRVACFACGG KLSNWEPKDN AMSEHLRHFP KCPFIENQLQ
DTSRYTVSNL SMQTHAARFK TFFNWPSSVL VNPEQLASAG FYYVGNSSDV KCFCCDGLR
CWESGDDPWV QHAKWFPRCE YLIRIKGQEF IRQVQASYPH LLEQLLSTSD SPGDENAESS
IIHFEPGEDH SEDAIMMNTF VINAAVEMGF SRSLVKQTVQ RKILATGENY RLVNDLVLDL
LNAEDEIREE ERERATEEKE SNDLLLIRKN RMALFQHLTC VIPILDSLLT AGIINEQEHD
VIKQKTQTSI QARELIDTIL VKGNIAATVF RNSLQEAQAV LYEHLEFVQD IKYIPTEDVS
DLPVEEQLRR LQEERTCKVC MDKEVSIVFI PCGHLVVCKD CAPSLRKCPI CRSTIKGTVR
TFLS
```

Figure 22

Figure 23

Name: BIR3Gen Bank Accession No. ACC41942.1SEQ ID No 23

SEQUENCE 618 AA

```
MHKTASQRLF PGPSYQNIKS IMEDSTILSD WTNSNKQKMK YDFSCELYRM STYSTFPAGV
PVSESLARA GFYYTGVNDK VKCFCCGLML DNWKLGDSPi QKHKQLYPSC SFIQNLVSAS
LGSTSKNTSP MRNSFAHSLs PTLHSSLFS GSYSSLSPNP LNSRAVEDIS SSRNTPYSYA
MSTEEARFLT YHMWPLTFLS PSELARAGFY YIGPGDRVAC FACGGKLSNW EPKDDAMSEH
RRHFPNCPFL ENSLETLRFS ISNLSMQTHA ARMRTFMYWP SSVPVQPEQL ASAGFYYVGR
NDDVKCFCCD GGLRCWESGD DPWVEHAKWF PRCEFLIRMK GQEFVDEIQG RYPHLLLEQLL
STSDTTGEEN ADPPIIHFGP GESSSEDAVM MNTPVVKSAL EMGFNRDLVK QTVQSKILTT
GENYKTVNDI VSALLNAEDE KREEEKEKQA EEMASDDLsL IRKNRMALFQ QLTCVLPILD
NLLKANVINK QEHDIIKQKT QIPLQARELI DTILVKGNAa ANIFKNCLKE IDSTLYKNLF
VDKNMKYIPT EDVSGLSLEE QLRLRQEERT CKVCMDKEVS VVFIPCGHLV VCQECAPSLR
KCPICRGIK GTVRTFLS
```

Figure 23

Figure 24

Name: CBLGen Bank Accession No. CAA40393.1SEQ ID No 24

SEQUENCE 906 AA

```
MAGNVKKSSG AGGGTGSGGS GSGGLIGLMK DAFQPHHHHH HHLSPHPPGT VDKKMVEKCW
KLMDKVVRLC QNPKLALKNS PPYILDLLPD TYQHLRTILS RYEGKMETLG ENEYFRVFME
NLMKTKQTI SLFKEGKERM YEENSQPRRN LTKLSLIFSH MLAEKLGIFP SGLFQGDTRF
ITKADAAEFW RKAFGEKTIV PWKSFRQALH EVHPISSGLE AMALKSTIDL TCNDYISVFE
FDIFTRLFQP WSSLLRNWNS LAVTHPGYMA FLTYDEVKAR LQKFIHKPGS YIFRLSCTRL
GQWAIGYVTA DGNILQTIPH NKPLFQALID GFREGFYLP DGRNQNPDLT GLCEPTQDH
IKVTQEYEL YCEMGSTFQL CKICAENDKD VKIEPCGHLM CTSCLTSWQE SEGQGCPCFR
CEIKGTEPIV VDPFDPRGSG SLLRQGAEGA PSPNYDDDDD ERADDTLFMM KELAGAKVER
PPSPFMAPQ ASLPPVPPRL DLLPQRVCVP SSASALGTAS KAASGSLHKD KPLPVPPTLR
DLPPPPPPDR PYSVGAESRP QRRPLPCTPG DCPSRDKLPP VPSSRLGDSW LPRPIPKVPV
SAPSSSDPWT GRELTNRHSL PPSLPSQMEP RPDVPRLGST FSLDTSMSMN SSPLVGPECB
HPKIKPSSA NAIYSLAARP LPVVKLPPEGE QCEGEEDTEY MTPSSRPLRP LDTSQSSRAC
DCDQQIDSCY YEAMYNISQ APSTTESSTF GEGNLAAHA NTGPESSENE DDGYDVPKPP
VPAVLARRTL SDISNASSSF GWLSLDGDPT TNVTEGSQVP ERPPKPPPRR INSERTKAGSC
QQSGGPAASA ATASPQLSSE IENLMSQGYG YQDIQKALVI AQNNIEMAKN ILREFVSISS
PAHVAT
```

Figure 24

Figure 25

Name: PMLGen Bank Accession No. AAA60388.1SEQ ID No 25

SEQUENCE 882 AA

```
MEPAPARSPR PQQDPARPQE PTMPPPETPS EGRQPSPPSPS PTERAPASEE EFQFLRCQQC
QAEAKCPKLL PCLHTLCSGC LEASGMQCPI CQAPWPLGAD TPALDNVFFE SLQRRLSVYR
QIVDAQAVCT RCKESADFWC FECEQLLCAK CFEAHQWFLK HEARPLAELR NQSVREFLDG
TRKTNNIFCS NPNHRTPTLT SIYCRGCSKP LCCSCALLDS SHSELKCDIS AEIQQRQEEL
DAMTQALQEQ DSAFGAVHAQ MHAAVGQLGR ARAETEELIR ERVRQVVAHV RAQERELLEA
VDARYQRDYE EMASRLGRLD AVLQRI RTGS ALVQRMKCYA SDQEVLDMHG FLRQALCRLR
QEEPQSLQAA VRTDGFDEFK VRLQDLSSCI TQKDAAVSK KASPEAASTP RDPIDVDLPE
EAERVKAQVQ ALGLAEAQPM AVVQSVPGAH PVPVYAFSIK GPSYGEDVSN TTTAQKRKCS
QTQCPRKVIK MESEEGKEAR LARSSPEQPR PSTSKAVSPP HLDGPPSPRS PVIGSEVFLP
NSNHVASGAG EAEERVVIS SSEDSDAENS SSRELDSSSS ESSDLQLEGP STLRLVDENL
ADPQAEDRPL VFFDLKIDNE TQKISQLAAV NRESKFRVVI QPEALFSIYS KAVSLEVGLQ
HFLSFLSSMR RPILACYKLW GPGLPNFFRA LEDINRLWEF QEAI SGFLAA LPLIRERVP
ASSFKLKNLA QTYLARNMSE RSAMAAVLAM RDLCLLLEVS PGPQLAQHVY PFSSLQCFAS
LQPLVQAAVL PRAEARLLAL HNVSMELLS AHRRDRQGGL KKYSRYLSLQ TTTLPPAQA
FNLQALGTYF EGLLEGPALA RAEGVSTPLA GRGLAERASQ QS
```

Figure 25

Figure 26

Name: RAG1Gen Bank Accession No. AAA60248.1SEQ ID No 26

SEQUENCE 1043 AA

```
MAASFPPTLG LSSAPDEIQH PHIKFSEWKF KLFVRVSFEK TPPEAQKEKK DSFEGKPSLE
QSPAVLDKAD GQKPVPTQPL LKAHPKFSKK FHDNEKARGK AIHQANLRHL CRICGNSFRA
DEHNRrypVH GPVDGKTLGL LRKKEKRATS WPDLIKVFR IDVKADVDSI HPTEFCHNCW
SIMHRKFSSA PCEVYFPRNV TMEWHPHTPS CDICNTARRG LKRKSLQPNL QLSKKLKTVL
DQARQARQRK RRAQARISSK DVMKKIANCS KIHLS TKLLA VDFPEHFVKS ISCQICEHIL
ADPVTENCKH VFCRVCILRC LKVMGSYCPS CRYPCFPTDL ESPVKSFLSV LNSLMVKCPA
KECNEEVSLK KYNHHISSHK ESKEIFVHIN KGGPRQHLL SLTRRAQKHR LRELKLQVKA
FADKEEGGDV KVCMTLFL ALRARNEHRQ ADELEAIMQG KGSGLPQAVC LAIRVNTFLS
CSQYHKMYRT VKAITGRQIF QPLHALRNAE KVL LPGYHHF EWQPPLKNVS SSTDVGIIDG
LSGLSSSVDD YPVDTIKRF RYDSALVSAL MDMEEDILEG MRSQDLDDYL NGPFTVVVKE
SCDGMGDVSE KHGSGPVVPE KAVRFSFTIM KITIAHSSQN VKVFEEAKPN SELCCKPLCL
MLADESDHET LTAILSPLIA EREAMKSSEL MLELGGILRT FKFI FRGTGY DEKLVREVEG
LEASGSVYIC TLCDATRLA SQNLVFHSIT RSHAENLERY EVWRSNPYHE SVEELRDRVK
GVSAPFIET VPSIDALHCD IGNAAEFYKI FQLEIGEVYK NPNASKEERK RWQATLDKHL
RKKMNLKPIM RMNGNFARKL MTKETVDAVC ELIPSEERHE ALRELMDLYL KMKPVWRSSC
PAKECPESLC QYSFNSQRFA ELLSTKFKYR YEGKITNYFH KTLAHVPEII ERDGSIGAWA
SEGNESGNKL FRRFRKMNR QSKCYEMEDV LKHHWLYTSK YLQKFMNAHN ALKTSGFTMN
PQASLGDP LG IEDSLESQDS MEF
```

Figure 26

Figure 27

Name: RAPSYNGen Bank Accession No. CAA83954.1SEQ ID No 27

SEQUENCE 305 AA

MASRKENAKS ANRVLRI SQL DALELNKALE QLVWSQFTQC FHGFKPGLLA RFEPEVKACL
WVFLWRFTIY SKNATVGQSV LNIKYKNDFS PNLRYQPPSK NQKIWYAVCT IGGRWLEERC
YDLFRNHHLA SFGKVKQCVN FVIGLLKLGG LINFLIFLQR GK FATLTERL LGIHSVFCKP
QNIREVGFEY MNRELLWHGF AEFLIFLLPL INVQKLKAKL SSWCIPLTGA PNSDNTLATS
GKECALCGER PTMPHTIGCE HIFCYFCAKS SFLFDVYFTC PKCGTEVHSL QPLKSGIEMS
EVNPL

Figure 27

Figure 28

Name: TIFI-ALPHAGen Bank Accession No. AAB63585.1SEQ ID No 28

SEQUENCE 1050 AA

```
MEVAVEKAVA AAAAAAAS GGPAAAPSGE NEAESRQGPD SERGGEAARL NLLDTCVCH
QNIQSRAPKL LPCLHSFCQR CLPAPQRYLM LPAPMLGSAE TPPPVPAPGS PVSGSSPFAT
QVGVIKCPVC SQECAERHII DNFFVKDTTE VPSSTVEKSN QVCTSCEDNA EANGFCVECV
EWLCKTCIRA HQRVKFTKDH TVRQKEEVSP EAVGVTSQRP VFCPFHKKEQ LKLYCETCDK
LTCRDCQLE HKEHRYQFIE EAFQNKQVII DTLITKLMEK TKYIKFTGNQ IQNRIIEVNQ
NQKQVEQDIK VAIFTLMVEI NKKGKALLHQ LESLAKDHRM KLMQQQQEVA GLSKQLEHVM
HFSKWAVSSG SSTALLYSKR LITYRLRHLL RARCASPVT NNTIQFHCDP SFWAQNIINL
GSLVIEDKES QPQMPKQNPV VEQNSQPPSG LSSNQLSKFP TQISLAQLRL QHMQQQVMAQ
RQQVQRRPAP VGLPNPRMQG PIQQPSISHQ QPPRLINFQ NHSPKPNGPV LPPHPQQLRY
PPNQNIQRQA IKPNPLQMAF LAQQAQKQWQ ISSGQGTPTT TNSTSTPSS PTITSAAGYD
GKAFGSPMID LSSPVGGSYN LPSLPDIDCS STIMLDNIVR KDTNIDHGQP RPPSNRTVQS
PNSSVPSPGL AGPVTMTSVH PPIRSPSASS VSGRSGSGSS SKPAGADSTH KVPVVMLEPI
RIKQENSGPP ENYDFPVVIV KQESDEESRP QNANYPRISL TSLLLNSSQS STSEETVLRN
DAPDSTGDQP GLHQDNSSNG KSEWLDPSQK SPLHVGETRK EDDPNEDWCA VCQNGGELLG
CEKCPKVFHL SCHVPTLTNF PSGEWICTFC RDLSPKEVEY DCDAPSHNSE KKKTEGLVKL
TPIDKRCER LLLFLYCHEM SLAFQDPVPL TVPDYKIIK NPMDLSTIKK RLQEDYSMYS
KPEDFVADFR LIFQNCAEFN EPDSEVANAG IKLENYFEEL LKNLYPEKRF PKPEFRNESE
DNKFSDSDSDD DVFQPRKKRL KSIEERQLLK
```

Figure 28

Figure 29

Name: TIFI-GAMMAGen Bank Accession No. AAD17259.1SEQ ID No 29

SEQUENCE 1127 AA

```
MAENKGGGEA ESGGGGSGSA PVTAGAAGPA AQEAEPLTA VLVEEEEEEG GRAGAEGGAA
GPDDGGVAAA SSGSAQAASS PAASVGTGVA GGAVSTPAPA PASAPAPGPS AGPPPGPPAS
LLDTCAVCQQ SLQSRREAEP KLLPCLHSFC LRCLPEPERQ LSVPIPGGSN GDIQQVG VIR
CPVCRQECRQ IDLVDNYFVK DTSEAPSSSD EKSEQVCTSC EDNASAVGFC VECGEWLCKT
CIEAHQVRKF TKDHLIRKKE DVSESVGASG QRPVFCPVHK QEQLKLCFET CDRLTCRDCQ
LLEHKEHRYQ FLEEAFQNQK GAIENLLAKL LEKKNYVHFA ATQVQNRIKE VNETNKRVEQ
EIKVAIFTLI NEINKKGKSL LQQLENTKE RQMKLLQQQN DITGLSRQVK HVMNFTNWAI
ASGSSTALLY SKRLITFQLR HILKARCDPV PAANGAIRFH CDPTFWAKNV VNLGNLVIES
KPAPGYTPNV VVGQVPPGTN HISKTPGQIN LAQLRLQHMQ QQVYAQKHQQ LQQMRMQQPP
APVPTTTTTT QQHPRQAAPQ MLQQQPRLI SVQTMQRGNM NCGAFQAHQM RLAQNAARIP
GIPRHSGPQY SMMQPHLQRQ HSNPGHAGPF PVVSVHNTTI NPTSPPTATM ANANRGPTSP
SVTAIELIPS VTNPENLPSL PDIPPIQLED AGSSSLDNLL SRYISGSHLP PQPTSTMNPS
PGPSALSPGS SGLSNSHTPV RPPSTSSTGS RGSCGSSGRT AEKTSLSFKS DQVKVKQEPG
TEDEICSFSG GVKQEKTEG RRSACMLSSP ESSLTPPLST NLHLESELDA LASLENHVKT
EPADMNESCK QSGLSSLVNG KSPTRSLMHR SARIGGDGNN KDDDPNEDWC AVCQNGGDL
CCEKCPKV FH LTCHVPTLLS FPSGDWICTF CRDIGKPEVE YDCDNLQHSK KGKTAQGLSP
VDQRKCERLL LYLYCHELSI EFQEPVPASI PNYKIIKKP MDLSTVKKKL QKKHSQHYQI
PDDFVADVRL IFKNCERFNE MMKVQVYAD TQEINLKADS EVAQAGKAVA LYFEDKLTEI
YSDRTFAPLP EFEQEEDDGE VTEDSDEDFI QPRRRLKSD ERPVHIK
```

Figure 29

Figure 30

Name: TRAF2Gen Bank Accession No. AAA87706.1SEQ ID No 30

SEQUENCE 501 AA

MAAASVTPPG SLELLQPGFS KTLGKLEA KYLCSACRNV LRRPFQAQCG HRYCSFCLAS
ILSSGFQONCA ACVHEGIYEE GISILESSSA FPDNAARREV ESLPAVCPSD GCTWKGTLKE
YESCHEGRCP LMLTECPACK GLVRLGEKER HLEHECPERS LSCRHCRAPC CGADVKAHHE
VCPKFPLTCD GCGKKKIPRE KFQDHVKTCG KCRVPCRFA IGCLETVEGE KQGEHEVQWL
REHLAMLLSS VLEAKPLLGD QSHAGSELLQ RCESEKKTA TFENIVCVLN REVERVAMTA
EACSRQHRLD QDKIEALSSK VQQLERSIGL KDLAMADLEQ KVRPFQAQCG HRYCSFCLAS
ILRKLQEAVA GRIPAIFSPA FYTSRYGYKM CLRIYLNQDG TGRGTHLSLF FVVMKGPND
LLRWPFNQKV TMLLDQNNR EHVIDAFRDP VTSSSFQRPV NDMNIASGCP LFCPVSKMEA
KNSYVRDDAI FIKAIVDLTG L

Figure 30

Figure 31

Name: TRAF3Gen Bank Accession No. AAC50112.1SEQ ID No 31

SEQUENCE 568 AA

MESSKKMDSP GALQTNPPLK LHTDRSAGTP VVFPEQGGYK EKFKVTVEDK YKCEKCHLVL
CSPKQTECGH RFCESCMAAL LSSSSPKCTA CQESIVKDKV FKDNCKREI LALQIYCRNE
SRGCAEQLTL GHLLVHLKND CHFEELPCVR PDCKEKVLRK DLRDHVEKAC KYREATCSHC
KSQVPMIALQ KHEDTDCPCV VVSCPHKCSV QTLRSELSA HLSECVNAPS TCSFKRYGCV
FQGTNQQIKA HEASSAVQHV NLLKEWSNSL EKKVSLQNE SVEKNKSIQS LHNQICSF EI
EIERQKEMLR NNEKILHLQ RVIDSQAEKL KELDKAIRPF RQNWEEADSM KSSVESLQNR
VTELESVDKS AGQVARNTGL LESQLSRHDQ MLSVHDIRLA DMDLRFQVLE TASYNGVLIW
KIRDYKRRKQ EAVMGKTLNL YSQPFYTGYP GYKMCARVYL NGDGMGKGTH LSLFFVIMRG
EYDALLPWWF KQKVTLMMLD QGSSRRHLGD AFKDPNSSL FKKPTGEMNI ASGCPVFVAQ
TVLENGTYIK DDTIFIKVIV DTSDLDPD

Figure 31

Figure 32

Name: HIP116Gen Bank Accession No. AAA67436.1SEQ ID No 32

SEQUENCE 1009 AA

```
MSWMFKRDPV WKYLQTVQYG VHGNFPRLSY PTFFPRFEFQ DVIPPDDFLT SDEEVDSVLF
GSLRGHVVGL RYYTGVVNNN EMVALQRDPN NPYDKNAIKV NNVNGNQVGH LKKELAGALA
YIMDNKLAQI EGVVPFGANN AFTMPLHMTF WGKEENRKAV SDQLKKHGFK LGPAPKTLGF
NLESGWGSGR AGPSYSMPVH AAVQMTTEQL KTEFDKLFED LKEDDKTHEM EPAAELETPL
LPHQKQALAW MVSRENSKEL PPFWEQRNDL YYNTITNFSE KDRPENVHGG ILADDMGLGK
TLTAIAVILT NFHDGRPLPI ERVKKNLLKK EYNVNDDSMK LGGNNTSEKA DGLSKDASRC
SEQPSISDIK ESKSKFRMSEL STSRPKRRKT AVQYIESSDS EEIETSELPQ KMKGKLNQVQ
SETKGRAKAG SSKVIEDVAF ACALTSSVPT TKKKMLKKGA CAVEGSKKT D VEERPRITLI
ICPLSVLSNW IDQFGQHIKS DVHLNFYVYY GPDRIREPAL LSKQDIVLTT YNLTTHDYGT
KGDSPHSIR WLRVILDEGH AIRNPNAQQT KAVLDLESER RWVLTGTPIQ NSLKDLWSLL
SFLKLKPFID REWWHRTIQR PVTMGDEGGL RRLQSLIKNI TLRRTKTSKI KGKPVLELPE
RKVFIQHITL SDEERKIYQS VKNEGRATIG RYFNEGTVLA HYADVLGLLL RLRQICCHTY
LLTNAVSSNG PSGNDTPEEL RKKLIRKMKL ILSSGSDEEC AICLDSLTPV VITHCAHVFC
KPCICQVIQN EQPHAKCPLC RNDIHEDNLL ECPPEELARD SEKKSDMEWT SSSKINALMH
ALTDLRKKNP NIKSLVVSQF TTFLSLIEIP LKASGFVFTR LDGSMQAKKR VESIQCQNT
EAGSPTIMLL SLKAGGVGLN LSAASRVFLM DPAWNPAED QCFDRCHRLG QKQEVITTKF
IVKDSVEENM LKIQNKREL AAGAFGTKKP NADEMKQAKI NEIRTLIDL
```

Figure 32

Figure 33

Name: STAF50Gen Bank Accession No. CAA57684.1SEQ ID No 33

SEQUENCE 442 AA

```
MDFSVKVDIE KEVTCPICLE LLTEPLSLDC GHSFCQACIT AKIKESVIIS RGESSCPVCQ
TRFQPGNLRP NRHLANIVER VKEVKMSPQE GQKRDVCEHH GKKLQIFCKE DGKVICWVCE
LSQEHQGHQT FRINEVVKEC QEKLQVALQR LIKEDQEAEK LEDDIRQERT AWKIERQKIL
KGFNEMRVIL DNEEQRELQK LEEGEVNVLD NLAAATDQLV QQRQDASTLI SDLQRRLTGS
SVEMLQDVID VMKRSESWTL KKP KSVSKKL KSVFRVPDLS GMLQVLKELT DVQYYWVDVM
LNPGSATS NV AISVDQRQVK TVRTCTFKNS NPCDFSAGV FGCQYFSSGK YYWEVDVSGK
IAWILGVHSK ISSLNKRKSS GFAFDPSVNY SKVYSRYPQ YGYWVIGLQN TCEYNAFEDS
SSSDPKVLTL FMAVLPVVLG FS
```

Figure 33

Figure 34

Name: HLTf-1Gen Bank Accession No. CAA86571.1SEQ ID No 34

SEQUENCE 1009 AA; 113914 MW; B4ACD86A6844CC5F CRC64;

MSWMFKRDPV WKYLQTVQYG VHGNTFRLSY PTFFLRFEFQ DVIPDDFLT SDEEVDSVLF
GSLRGHVVL RYYTGVVNNN EMVALQRDPN NPYDKNAIKV NNVNGNQVGH LKKELAGALA
YIMDNKLAQI EGVVPFGANN AFTMPLHMTF WGKEENRKAV SDQLKKHGFK LGPAPKTLGF
NLESGWGSGR AGPSYSMPVH AAVQMTTEQL KTEFDKLFED LKEDDKTHEM EPAEAIETPL
LPHQKQALAW MVSRENSKEL PPFWEQRNDL YNTITNFSE KDRPENVHGG ILADDMGLGK
TLTAIAVILT NFHDGRPLPI ERVKKNLLKK EYNVNDGSMK LGGNNTSEKA DGLSKDASRC
SEQPSISDIK EKSKFRMSEL SSSRPKRRKT AVQYIESSDS EEIETSELPQ KMKGKLKNVQ
SETKGRKAG SSKVIEDVAF ACALTSSVPT TKKKMLKKGA CAVEGSKKTD VEERPRTTLI
ICPLSVLSNW IDQFGQHIKS DVHLNFYVYY GPDRIREPAL LSKQDIVLTT YNLTIDYGT
KGDSPHLSIR WLRVILDEGH AIRNPNAQQT KAVLDLESER RWVLTGTPIQ NSLKDLWSLL
SFLKLKPFID REWWHRTIQR PVTMGDEGGL RRLQSLIKNI TLRRTKTSKI KGKPVLELPE
RKVFIQHITL SDEERKIYQS VKNEGRATIG RYFNEGTVLA HYADVLGLLL RLRQICCHTY
LLTNAVSSNG PSGNDTPEEL RKKLIRKMKL ILSSGSDEEC AICLDSLTPV VITHCAHVFC
KPCICQVIQN EQPHAKCPLC RNDIHEDNLL ECPPEELARD SEKKSDMEWT SSSKINALMH
ALTDLRKKNP NIKSLVVSQF TTFLSLIEIP LKASGFVFTR LDGSMAQKKR VESIQCQNT
EAGSPTIMLL SLRAGGVGLN LSAASRVFLM DPAWNPAAED QCFDRCHRLG QKQEVITTKF
IVKDSVEENM LKIQNKKREL AAGAFGTKKP NADEMKQAKI NEIRTLIDL

//

Figure 34

Figure 35

Name: BIR4Gen Bank Accession No. AAC50373.1SEQ ID No 35

SEQUENCE 497 AA

```
MTFNSFEGSK TCVPADINKE EEFVEEFNRL KTFANFPSGS PVSASTLARA GFLYTGEEDT
VRCFSCHAAV DRWQYGDSAV GRHRKVSPNC RFINGFYLEN SATQSTNSGI QNGQYKVENY
LGSRDHFALD RPSETHADYL LRTGQVVDIS DTIYPRNPAM YSEEARLKSF QNWPDYAHLT
PRELASAGLY YTGIGDQVQC FCCGGKLKNW EPCDRAWSEH RRHFPNCFFV LGRNLNIRSE
SDAVSSDRNF PNSTNLPRNP SMADYEARIF TFGTWIYSVN KEQLARAGFY ALGEGDKVKC
FHC GGGLTDW KPSEDPWEQH AKWYPGCKYL LEQKGQEYIN NIHLTHSLEE CLVRTTEKTP
SLTRRIDDTI FQNPMVQEAI RMGFSFKDIK KIMEEKIQIS GSNYKSLEVL VADLVNAQKD
SMQDESSQTS LQKEISTEEQ LRRLOEEKLC KICMDRNIAI VFVPCGHLVT CKQCABAVDK
CPMCYTVITF KQKIFMS
```

Figure 35

Figure 36

Name: ANDENOVIRUS 5E1A- Gen Bank Accession No. CAA60052.1
3INDING PROTEIN

SEQ ID No 36

SEQUENCE 562 AA

```
MSRVHGMHPK ETTRQLSLAV KDGLIVETLT VGCKGSKAGI EQEGYWLPD EIDWETENHD
WYCFEHLPG EVLICDLCPV VYHSKCLSDE FRLRDSSSPW QCPVCRSIKK KNTNKQEMGT
YLRFIVSRMK ERAIDLNKKG KDNKHPMYRR LVHSAVDVPT IQEKVNEGKY RSYEEFKADA
QLLLHNTVIF YGADSEQADI ARMLYKDTCH ELDELQLCKN CFYLSNARPD NWFCYPCIPN
HELWVAKMKG FGFWPAKVMQ KEDNQVDVRF FGHHHQRAWI PSENIQDITV NIHRLHVKRS
MGWKKACDEL ELHQRFLREG RFWKSKNEDR GEEEAESSIS STSNEQLKVT QEPRAKKGRR
NQSVEPKKEE PEPETEAVSS SQEIPTMPQP IEKVSVSTQT KKLSASSPRM LHRSTQTTND
GVCQSMCHDK YTKIFNDFKD RMKSDHKRET ERVVREALEK LRSEMEEEEKR QAVNKAVANM
QGEMDRKCKQ VKEKCKEEFV EEIKKLATQH KQLISQTKKK QWCYNCEEEA MYHCCWNTSY
CSIKCQEHW HAEHKRTCRR KR
```

Figure 36

Figure 37

Name: HT2AGen Bank Accession No. AAA86474.1SEQ ID No 37

SEQUENCE 653 AA

```
MAAAAASHLN LDALREVLEC PICMESITEE QLRPKLLHCG HTICRQCLEK LlassingVR
CPFCSKITRI TSITQLTDNL TVLKIIDTAG LSEAVGLLMC RSCGRRLP RQ FCRSCGLVLC
EPCREADHQP PGHCTLPVKE AAEEERRRDFG EKLTRLRELM GELQRRKAAL EGVSKDLQAR
YKAVLQ EYGH EERRVQDELA RSRKFFTGSL AEVEKSNSQV VEEQSYLLNI AEVQAVSRCD
YFLAKIKQAD VALLEETADE EEP ELTASLP RELTLQDVEL LKVGHVGPLQ IGQAVKKPRT
VNVEDSWAME ATASAASTSV TFREMDMSPE EVVASPRASP AKQRGPEAAS NIQQCLFLKK
MGAKGSTPGM FNL PVS LYVT SQGEVLVADR GNYRIQVFTR KGFLKEIRRS PSGIDSFVLS
FLGADLPNLT PLSVAMNCQG LIGV TDSYDN SLKVYTLDGH CVACHRSQLS KPWGITALPS
GQFVVTDVEG GKLWCFTVDR GSGVVKYSCL CSAVRPKFVT CDAEGTVYFT QGLGLNLENR
QNEHHLEGGF SIGSVGPDGQ LGRQISHFFS ENEDFRCIAG MCV DARGDLI VADSSRKEIL
HFPKGGGYSV LIREGLTCPV GIALTPKGQL LVLD CWDHCI KIYSYHLRRY STP
```

//

Figure 37

Figure 38

Name: TRAF3Gen Bank Accession No. AAC50112.1SEQ ID No 38

SEQUENCE 568 AA

MESSKKMDSP GALQTNPPLK LHTDRSAGTP VVVPEQGGYK EKFVKTVEDK YKCEKCHLV
CSPKQTECGH RFCESCMAAL LSSSSPKCTA CQESIVKDKV FKDNCKCKREI LALQIYCRNE
SRGCAEQTLT GHLLVHLKND CHFEELPCVR PDCKEKVLRL DLRDHVEKAC KYREATCSHC
KSQVPMIALQ KHEDTDCPCV VVSCPHKCSV QTLRLSELSA HLSECVNAPS TCSFKRYGCV
FQGTNQIKA HEASSAVQHV NLLKEWSNSL EKKVSLQNE SVEKNKSIQS LHNQICSFEI
EIERQKEMLR NNEKILHLQ RVIDSQAELK KELDKAIRPF RQNWEEADSM KSSVESLQNR
VTELESVDKS AGQVARNTGL LESQLSRHDQ MLSVHDIRLA DMDLRFQVLE TASYNGVLIW
KIRDYKRRKQ EAVMGKTLNL YSQPFYTGYP GYKMCARVYL NGDGMGKGTH LSLFFVIMRG
EYDALLPWP KQKVTLMMLD QGSSRRHLGD AFKPDPNSSS FKKPTGEMNI ASGCPVFVAQ
TVLENGTYIK DDTIFIKVIV DTSDLDPD

Figure 38

Figure 39: Human PRT3 Coding Sequence (SEQ ID NO:40)

ATGGATGAATCAGCCTTGTGATCTTTTGGAGTGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGA
AGGTCTTGCCCTTGCCAGCATACGTTTTGCAAGCGATGTTTGCTGGGGATCGTAGGTTCTCGAAATGAACT
CAGATGTCCCGAGTGCAGGACTCTTGTTGGCTCGGGTGTGAGGAGCTTCCAGTAACATCTTGCTGGTC
AGACTTCTGGATGGCATCAAACAGAGGCCTTGAAACCTGGTCTGGTGGGGGAAGTGGGACCAACTGCA
CAAATGCATTAAGGTCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAAGATCTGCAGAGCTCCAGGG
CGGACAGCAGCCTCGGGTGCAATCCTGGAGCCCCCAGTGAGGGGTATACCTCAGTTACCATGTGCCAAA
GCGTTATACAACATATGAAGGAAAAGAGCCTGGAGACCTTAAATTCAGCAAAGGCGACATCATCATTTTGC
GAAGACAAGTGGATGAAAATTGGTACCATGGGGAAGTCAATGGAATCCATGGCTTTTTCCCCACCAACTT
TGTGCAGATTATTAAACCGTTACCTCAGCCCCCACCCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAA
GACAAGGAAGCAGACAAAGATTGCCTTCCATTGCAAAGGATGATGTTCTGACTGTGATCCGAAGAGTGG
ATGAAAACCTGGGCTGAAGGAATGCTGGCAGACAAAATAGGAATATTTCCAATTTCATATGTTGAGTTTAA
CTCGGCTGCTAAGCAGCTGATAGAATGGGATAAGCCTCCTGTGCCAGGAGTTGATGCTGGAGAATGTTCC
TCGGCAGCAGCCAGAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACACCAAAAAGCGGCACT
CCTTCACTTCCCTCACTATGGCCAACAAGTCTCCAGGCATCCAGAACCGCCACTCCATGGAGATCAG
CCCCCTGTCTCATCAGCTCCAGCAACCCCACTGCTGCTGCACGGATCAGCGAGCTGTCTGGGCTCTCC
TGCACTGCCCCCTTCTCAGGTTTCATATAAGTACCACCGGTTAATTGTGACCCCCGCCCCAAGCAGCCCAG
TGACAACTGGCCCCCTCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACTTTGAATCC
TCCTCTTCCACCACCCCTCTCCTGGCTGCCACTGTCTTGCCTCCACACCACCAGGCGCCACCGCCGCC
GCTGCTGCTGCTGGAATGGGACCGAGGCCCATGGCAGGATCCACTGACCAGATTGCACATTTACGGCCGC
AGACTCGCCCCAGTGTGTATGTTGCTATATATCCATACACTCCTCGGAAAGAGGATGAACTAGAGCTGAG
AAAAGGGGAGATGTTTTTAGTGTGTTGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATGCATACC
AGCAAGATAGGGGTTTTCCCTGGCAATTATGTGGCACCAGTCACAAGGGCGGTGACAAATGCTTCCCAAG
CTAAAGTCCCTATGTCTACAGCTGGCCAGACAAGTCGGGGAGTGACCATGGTCAGTCTTCCACGGCAGG
AGGGCCTGCCCAGAAGCTCCAGGGAAATGGCGTGGCTGGGAGTCCCAGTGTGTCCTCCGACGCTGTGGTA
TCAGCAGCTCACATCCAGACAAGTCTCAGGCTAAGGTCTTGTGTCACATGACGGGGCAAATGACAGTCA
ACCAGGCCCGCAATGCTGTGAGGACAGTTGCAGCGCACAAACCAGGAACGCCCCACGGCAGCAGTGACACC
CATCCAGGTACAGAATGCCGCCGGCCTCAGCCCTGCATCTGTGGGCTGTCCCATCACTCGCTGGCCTCC
CCACAACCTGCGCCTCTGATGCCAGGCTCAGCCACGCACACTGCTGCCATCAGTATCAGTCGAGCCAGTG
CCCCCTGTGGCCTGTGCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGCTTCTCTGGAGGCTGA
GCCCAGTGGCCGGATAGTGACCGTTCTCCCTGGACTCCCCACATCTCCTGACAGTGCTTCATCAGCTTGT
GGGAACAGTTTCAGCAACCAAAACCAGACAAGGATAGCAAAAAGAAAAAAGGGTTTGTGTAAGTTGCTTT
CTGGCGCCTCCACTAAACGGAAGCCCCGCGTGTCTCTCCAGCATCGCCACCCTAGAAGTGGAGCTGGG
CAGTGCAGAGCTTCTCTCCAGGGAGCGGTGGGGCCGAACTGCCACCAGGAGGTGGCCATGGCAGGGCA
GGCTCCTGCCCTGTGGACGGGACGGACCGGTACGACTGCAGTGGCAGGAGCAGCCCTGGCCCAGGATG
CTTTTCATAGGAAGGCAAGTTCCCTGGACTCCGCAGTTCATCGCTCCACCTCCTCGCCAGGCCTGTTC
CTCCCTGGGTCCTGTCTGAATGAGTCTAGACCTGTCGTTTGTGAAAGGCACAGGGTGGTGGTTTCTAT
CCTCCTCAGAGTGAGGCAGAACTTGAACCTAAAGAAGGAGATATTGTGTTTGTTCATAAAAAACGAGAGG
ATGGCTGGTTCAAAGGCACATTACAACGTAATGGGAAAACCTGGCCTTTTCCAGGAAGCTTTGTGAAAA
CATATGA

Figure 40: Human PRT3 cDNA Sequence (SEQ ID NO:41)

CTGAGAGACACTGCGAGCGGCGAGCGCGGTGGGGCCGCATCTGCATCAGCCGCCGAGCCGCTGCGGGGG
CGCGAACAAAGAGGAGGAGCCGAGGCGCGAGAGCAAAGTCTGAAATGGATGTTACATGAGTCATTTTAAG
GGATGCACACAACATATGAACATTTCTGAAGATTTTTTCTCAGTAAAGTAGATAAAGATGGATGAATCAGC
CTTGTGGATCTTTTGGAGTGTCCGGTGTGTCTAGAGCGCCTTGATGCTTCTGCGAAGGTCTTGCCCTTG
CAGCATACGTTTTGCAAGCGATGTTTGTGCGGGATCGTAGGTTCTCGAAATGAACTCAGATGTCCCGAGT
GCAGACTCTTGTGGCTCGGGTGTGAGGAGCTTCCCAGTAACATCTTGCTGGTCAGACTTCTGGATGG
CATCAAACAGAGGCCCTTGAAACCTGGTCTTGGTGGGGGAAGTGGGACCAACTGCACAAATGCATTAAGG
TCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAAGATCTGCAGAGCTCCAGGGCGGACAGCAGCCTC
GGGTGCAATCCTGGAGCCCCCAGTGAGGGGTATACCTCAGTTACCATGTGCCAAAGCGTTATACAACATA
TGAAGGAAAAGAGCCTGGAGACCTTAAATTCAGCAAAGCGACATCATCATTTTGCGAAGACAAGTGGAT
GAAAATTGGTACCATGGGGAAGTCAATGGAATCCATGGCTTTTTCCCCACCAACTTTGTGCAGATTATTA
AACCCTTACCTCAGCCCCACCTCAGTGCAAAGCACTTTATGACTTTGAAGTGAAAGACAAGGAAGCAGA
CAAAGATTGCCTTCCATTTGCAAAGGATGATGTTCTGACTGTGATCCGAAGAGTGGATGAAAACCTGGGCT
GAAGGAATGCTGGCAGACAAAATAGGAATATTTCCAATTTTCATATGTTGAGTTTAACTCGGCTGCTAAGC
AGCTGATAGAATGGGATAAGCCTCCTGTGCCAGGAGTTGATGCTGGAGAATGTTCTCGGCAGCAGCCCA
GAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACACCAAAAAGCGGCACTCCTTCACTTCCCTC
ACTATGGCCAACAAGTCTCCTCCAGGCATCCAGAACCGCCACTCCATGGAGATCAGCCCCCCTGTCCCTCA
TCAGCTCCAGCAACCCCACTGCTGCTGCACGGATCAGCGAGCTGTCTGGGCTCTCCTGCAGTGCCTTCTC
TCAGGTTTCATATAAGTACCACCGGGTTAATTGTGACCCCCGCCCAAGCAGCCAGTGACAACCTGGCCCC
TCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACCTTTGAATCCTCCTCTTCCACCAC
CCCCCTCTCCTGGCTGCCACTGTCTTGCCTCCACACCACCGGCGCCACCGCCGCGCTGCTGCTGCTGG
AATGGGACCGAGGCCCATGGCAGGATCCACTGACCAGATTGCACATTTACGGCCGCAGACTCGCCCCAGT
GTGTATGTTGCTATATATCCATACACTCCTCGGAAAAGAGGATGAACTAGAGCTGAGAAAAGGGGAGATGT
TTTTAGTGTGTTGAGCGCTGCCAGGATGGCTGGTTCAAAGGGACATCCATGCATACCAGCAAGATAGGGGT
TTTCCCTGGCAATTATGTGGCACCAGTCAACAAGGGCGGTGACAAATGCTTCCCAAGCTAAAGTCCCTATG
TCTACAGCTGGCCAGACAAGTCCGGGAGTGACCATGGTCACTGACGGGCAATGACAGTCAACCAGGCCGCAAT
AGCTCCAGGAAATGGCGTGGCTGGGAGTCCAGTGTGTGCTCCCGCAGCTGTGGTATCAGCAGCTCACAT
CCAGACAAGTCTCAGGCTAAGGTCTTGTGCATGACGGGGCAATGACAGTCAACCAGGCCGCAAT
GCTGTGAGGACAGTTGCAGCGCACAAACGGAACGCCCCACGGCAGCAGTGACACCCATCCAGGTACAGA
ATGCCGCGGCGCTCAGCCCTGCATCTGTGGGCTGTCCCATCACTCGCTGGGCTCCCCACAACCTGCGGCC
TCTGATGCCAGGCTCAGCCACGCACACTGCTGCCATCAGTATCAGTCGAGCCAGTGCCCCCTCTGGCCTGT
GCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGCTTCTCTGGAGGCTGAGCCCACTGGCCGGA
TAGTGACCGTTCTCCCTGGACTCCCCACATCTCCTGACAGTGCTTCATCAGCTTGTGGGAACAGTTCAGC
AACCAAACCAGACAAGGATAGCAAAAAAGAAAAAAGGGTTTGTGTAAGTTGCTTTCTGGCGCTCCACT
AAACGGAAGCCCCGCGTGTCTCCTCCAGCATCGCCACCCCTAGAAGTGGAGCTGGGCAGTGACAGAGCTTC
CTCTCCAGGGAGCGGTGGGGCCCCGAAC TGCCACCAGGAGGTGGCCATGGCAGGGCAGGCTCCTGCCCTGT
GGACGGGGACGGACCGGTACGACTGCAGTGGCAGGAGCAGCCCTGGCCAGGATGCTTTTCATAGGAAG
GCAAGTTCCCTGGACTCCGCAGTTCCCATCGCTCCACCTCCTCGCCAGGCCTGTTCTCTCCTGGGTCTGT
TCTTGAATGAGTCTAGACCTGTGCTTGTGAAAGGCACAGGGTGGTGGTTTCTATCCTCCTCAGAGTGA
GGCAGAACTTGAACCTTAAAGAAGGAGATATTGTGTTTGTTCATAAAAAACGAGAGGATGGCTGGTTCAAA
GGCAGATTACAACGTAATGGGAAAACCTGGCCCTTTCCAGGAAGCTTTGTGGAAAACATATGAGGAGACT
GACACTGAAGAAGCTTAAAAATCACTTCACACAACAAAGTAGCACAAGCAGTTTAAACAGAAAGAGCACAT
TTGTGGACTTCCAGATGGTCAGGAGATGAGCAAAGGATTGGTATGTGACTCTGATGCCCCAGCACAGTTA
CCCCAGCGAGCAGAGTGAAGAAGATGTTTGTGTGGGTTTTGTAGTCTGGATTCCGATGTATAAGGTGTG
CCTTGTACTGTCTGATTTACTACACAGAGAACTTTTTTTTTTTTTTAAAGATATATGACTAAAATGGACA
ATTGTTTACAAGGCTTAACTAATTTATTTGCTTTTTTAAACTTGAACCTTTTCGTATAATAGATACGTTCT
TTGGATTATGATTTTAAAGAAATATTAATTTATGAAATGATAGGTAAGGAGAAGCTGGATTATCTCCTGT
TGAGAGCAAGAGATTCTTTTGACATAGAGTGAATGCATTTTCCCCTCTCCTCCTCCTGCTACCATTTAT
ATTTTGGGGTTATGTTTTGCTTCTTTAAGATAGAAATCCCAGTTCTCTAATTTGGTTTTCTTCTTTGGGA
AACCAAACATACAAATGAATCAGTATCAATTAGGGCTGGGGTAGAGAGACAGAACTTGAGAGAAGAGA
AGTTAGTGATTCCCTCTCTTTCTAGTTTGGTAGGAATCACCCCTGAAGACCTAGTCTCAATTTAATTGTG
TGGGTTTTTAATTTCTTAGAATGAAGTGAAGTGAACAAATGAGAAAGAAATACAGCACAAACCTTGAACAA
AATGTATTTAGAAATATATTTAGTTTATAGCAGAAGCAGCTCAATTTGTTTGGTTGGAAAGTAGGGGAAA
TTGAAGTTGTAGTCACTGTCTGAGAATGGCTATGAAGCGTCATTTACATTTTACCCCAACTGACCTGCA
TGCCAGGACACAAGTAAGTAAACATTTGTGAGATAGTGGTGGTAAGTGATGCACTCGTGTAAAGTCAAAGGC
TATAAGAAAACACTGTGAAAAGTTTCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
GGATTCCCACAGTAATATAGACTGTGCATGGTGTGTATATTTTATTTTATTTTATTTTATTTTATTTTATTTT
GTACTCAGAATTGACCAATTGAGGAGGTGTAATAAATAACAGTGTCTCTCTCTACCCCAAAGCCACTA

CTGACCAAGGTCTCTTCAGTGCACTCGCTCCCTCTCTGGCTAAGGCATGCATTAGCCACTACACAAGTCA
T TAGTGAAAGTGGTCTTTTATGTCTCTCCAGCAGACAGACATCAAGGATGAGTTAACCAGGAGACTACTC
CTGTGACTGTGGAGCTCTGGGAAGGCTTGGTGGGAGTGAATTTGCCACACCTTACAATTGTGGCAGGATC
CAGAAGAGCCTGTCTTTTATATCCATTCTTGATGTCAATTGGCCTCTCCCACCGATTCATTACGGTGC
CACGCAGTCATGGATCTGGGTAGTCCGGAACAAAAGGAGGGAAGACAGCCTGGTAATGAATAAGATCC
TTACCACAGTTTTCTCATGGGAAATACATAATAAACCTTTTCATCTTTTTTTTTTTCTTTAAGAATTAA
AACTGGGAAATAGAAACATGAACTGAAAAGTCTTGCAATGACAAGAGGTTTCATGGTCTTAAAAAGATAC
TTTATATGGTTGAAGATGAAATCATTCCTAAATTAACCTTTTTTTTTTAAAAAAAACAATGTATATTATGT
TCCTGTGTGTTGAATTTAAAAAAAAAAAAATACTTTACTTGGATATTCATGTAATATATAAAGGTTTGGTG
AAATGAACTTTAGTTAGGAAAAAGCTGGCATCAGCTTTTCATCTGTGTAAGTTGACACCAATGTGTCATAA
TATTCTTTATTTTGGGAAATTAGTGATTTTATAAAAAATTTAAAAAGAAAAAAGACTACTACAGGTAA
GATAATTTTTTTACCTGTCTTTTCTCCATATTTAAGCTATGTGATTGAAGTACCTCTGTTTCATAGTTTC
CTGGTATAAAGTTGGTTAAAAATTCATCTGTTAATAGATCATTAGGTAATATAATGTATGGGTTTTCTAT
TGGTTTTTTTGCAGACAGTAGAGGGAGATTTTGTAAACAAGGGCTTGTTACACAGTGATATGGTAATGATAA
AATTGCAATTTATCACTCCTTTTCATGTTAATAATTTGAGGACTGGATAAAAGGTTTCAAGATTAAATTT
TGATGTTCAAACCTTTGT

Figure 41: 5' cDNA fragment of human PRT3 (public gi:10432611; SEQ ID NO:42)

ctgagagacactgcgagcggcgagcgcgggtggggccgcacatctgcatcagccgcgcgagccgctgcggggc
cgcgaaacaaagaggaggagccgaggcgcgagagcaaaagtctgaaatggatgttacatgagtcattttaag
gatgcacacaactatgaacatttctgaagattttttctcagtaaagtagataaagatggatgaatcagcc
ttgttggatcttttggagtgctcgggtgtgtctagagcgccttgatgcttctggaaggtcttgcccttgcc
agcatacgttttgcgaagcgatgtttgctggggatcgtaggttctcgaaatgaactcagatgtcccgagt
caggactcttggtggctcgggtgtcgaggagcttcccagtaacatcttgctggtcagacttctggatggc
atcaaacagagggccttggaaacctgggtcctgggtgggggaagtgggaccaactgcacaaatgcattaagg
ctcagagcagcactgtggctaattgtagctcaaaagatctgcagagctcccagggcggacagcagcctcg
gggtgcaatcctggagccccccagtgaggggtatacctcagttaccatgtgccaaagcgttatacaactat
gaaggaaaagagcctggagaccttaaatcagcaaaggcgacatcatcattttggaagacaagtggatg
aaaattggtaccatggggaagtcaatggaatccatggctttttccccaccaactttgtgcagattattaa
accgttacctcagccccccacctcagtgcaaagcactttatgactttgaagtgaagacaaggaagcagac
aaagattgccttccatttgcaaaggatgatgttctgactgtgatccgaagagtggatgaaaactgggctg
aaggatgctggcagacaaaataggaatatttccaatttcatatgttgagtttaactcggctgctaagca
gctgataagaatgggataagcctcctgtgccaggagttgatgctggagaatgttcctcggcagcagccag
agcagcactgccccaaagcactccgacaccaagaagaacacccaaaagcgggcactccttcacttcctca
ctatggccaacaagtccctccaggcatcccagaaccggccactccatggagatcagccccctgtcctcat
cagctccagcaacccccactgctgctgcacggatcagcgagctgtctgggctctcctgcagtgcccttct
caggttcatataagtaccaccgggttaattgtgaccccgcccccaagcagcccagtgacaactggccccct
cgtttactttcccatcagatgttccctaccaagctgcccttggaaactttgaatcctcctcttccaccacc
ccctctcctggctgccactgtccttgctccacaccacccaggcgccacgcgcgcgcgctgctgctgga
atgggaccgaggcccatggcaggatccactgaccagattgcacatttacggccgcagactcgccccagtg
tgtatgttgctatatatccatacactcctcggaagaggatgaactagagctgagaaaaggggagatgtt
tttagtggttgagcgctgccaggatggctggttcaaaggacatccatgcataccagcaagataggggtt
ttccctggcaattatgtggcaccagtcacaagggcggtgacaaatgcttcccaagctaaagtccctatgt
ctacagctggccagacaagtccgggagtgaccatggtcagtccttccacggcaggagggcctgccagaa
gctccaggggaaatggcgctggctgggagtcacagtggttccccgcagctgtggtatcagcagctcacatc
cagacaagtcctcaggctaaggctcttggtgcacatgacggggcaaatgacagtcacccaggcccgcaatg
ctgtgaggacagttgcagcgcacaaccaggaacgccccacggcagcagtgacacccatccaggtacagaa
tgccgcggcctcagccctgcacatctgtgggcctgtcccatcactcgctggcctccccacaacctgcgcct
ctgatgccaggctcagccacgcacactgtgccatcagtatcagtcgagccagtgccctctggcctgtg
cagcagctgtccactgacttccccaaagcatcaccagtgcttctctggaggctgagcccagtgccggat
agtgaccgttctcctggactccccacatctcctgacagtgcttcatcagcttgtgggaacagttcagca
accaaaccagacaaggatagc

Figure 42: 3' mRNA fragment of hPOSH (public gi:7959248; SEQ ID NO:43)

at t t c a t a t g t t g a g t t t a a c t c g g c t g c t a a g c a g c t g a t a g a a t g g g a t a a g c c t c c t g t g c c a g g a g
t t g a t g c t g g a g a a t g t t c c t c g g c a g c a g c c c a g a g c a g c a c t g c c c c a a g c a c t c c g a c a c c a a g a a
g a a c a c c a a a a a g c g g c a c t c c t t c a c t t c c c t c a c t a t g g c c a c a a g t c c t c c c a g g c a t c c c a g a a c
c g c c a c t c c a t g g a g a t c a g c c c c c t g t c c t c a t c a g c t c c a g c a a c c c c a c t g c t g c t g c a c g g a t c a
g c g a g c t g t c t g g g c t c t c c t g c a g t g c c c c t t c t c a g g t t c a t a t a a g t a c c a c c g g g t t a a t t g t g a c
c c c g c c c c a a g c a g c c c a g t g a c a a c t g g c c c t c g t t t a c t t t c c c a t c a g a t g t t c c c t a c c a a g c t
g c c c t t g g a a c t t t g a a t c c t c c t c t t c c a c c a c c c c t c t c c t g g c t g c c a c t g t c c t t g c c t c c a c a c
c a c c a g g c g c c a c c g c g c t g c t g c t g c t g c t g g a a t g g g a c c g a g g c c c a t g g c a g g a t c c a c t g a c c a
g a t t g c a c a t t t a c g g c c g c a g a c t c g c c c a g t g t g a t g t t g c t a t a t a t c c a t a c a c t c c t c g g a a a
g a g g a t g a a c t a g a g c t g a g a a a g g g g a g a t g t t t t t a g t g t t g a g c g c t g c c a g g a t g g c t g g t t c a
a a g g g a c a t c c a t g c a t a c c a g c a a g a t a g g g g t t t t c c c t g g c a a t t a t g t g g c a c c a g t c a c a a g g g c
g g t g a c a a a t g c t t c c c a a g c t a a a g t c c c t a t g t c t a c a g c t g g c c a g a c a a g t c g g g g a g t g a c c a t g
g t c a g t c c t t c c a c g g c a g g a g g g c c t g c c c a g a a g c t c c a g g g a a a t g g c g t g g c t g g g a g t c c c a g t g
t t g t c c c c g c a g c t g t g g t a t c a g c a g c t c a c a t c c a g a c a a g t c c t c a g g c t a a g g t c t t g t t g c a c a t
g a c g g g g c a a a t g a c a g t c a a c c a g g c c c g c a a t g c t g t g a g g a c a g t t g c a g c g c a c a a c c a g g a a c g c
c c c a c g g c a g c a g t g a c a c c c a t c c a g g t a c a g a a t g c c g c c g g c c t c a g c c c t g c a t c t g t g g g c c t g t
c c c a t c a c t c g c t g g c c t c c c c a c a a c c t g c g c c t c t g a t g c c a g g c t c a g c c a c g c a c a c t g c t g c c a t
c a g t a t c a g t c g a g c a g t g c c c c t c t g g c c t g t g c a g c a g c t g c t c c a c t g a c t t c c c c a a g c a t c a c c
a g t g c t t c t c t g g a g g c t g a g c c a g t g g c c g g a t a g t g a c c g t t c c c c t g g a c t c c c c a c a t c t c c t g
a c a g t g c t t c a t c a g c t t g t g g g a a c a g t t c a g c a a c c a a a c c a g a c a a g g a t a g c a a a a a g a a a a a a
g g g t t t g t t g a a g t t g c t t t c t g g c g c c t c c a c t a a a c g g a a g c c c c g c g t g t c t c c t c c a g a t c g c c c
a c c c t a g a a g t g g a g c t g g g c a g t g c a g a g c t t c c t c t c c a g g g a g c g g t g g g g c c c g a a c t g c c a c c a g
g a g g t g g c c a t g g c a g g g c a g g c t c c t g c c c t g t g g a c g g g a c g g a c c g g t c a c g a c t g c a g t g g c a g g
a g c a g c c c t g g c c c a g g a t g c t t t t c a t a g g a a g g c a a g t t c c c t g g a c t c c g c a g t t c c c a t c g c t c c a
c c t c c t c g c c a g g c c t g t t c c t c c c t g g g t c c t g t c t t g a a t g a g t c t a g a c c t g t c g t t t g t g a a g g c
a c a g g g t g g t g g t t t c c t a t c c t c c t c a g a g t g a g g c a g a a c t t g a a c t t a a g a a g g a g a t a t t g t g t t
t g t t c a t a a a a a a c g a g a g g a t g g c t g g t t c a a a g g c a c a t t a c a a c g t a a t g g g a a a c t g g c c t t t t c
c c a g g a a g c t t t g t g g a a a c a t a t g a g g a g a c t g a c a c t g a a g a a g c t t a a a t c a c t t c a c a c a a c a a
a g t a g c a c a a a g c a g t t t a a c a g a a a g a g c a c a t t t g t g g a c t t c c a g a t g g t c a g g a g a t g a g c a a a g g
a t t g g t a t g t g a c t c t g a t g c c c a g c a c a g t t a c c c a g c g a g c a g a g t g a a g a a g a t g t t t g t g t g g g
t t t t t g t a g t c t g g a t t c g g a t g t a a a g g t g c c t t g t a c t g t c t g a t t t a c t a c a c a g a a a c t t t
t t t t t t t t t t a a g a t a t a t g a c t a a a a t g g a c a a t t g g t t t a c a a g g c t t a a c t a a t t t a t t t g c t t t t
t a a a c t t g a a c t t t t c g t a t a a t a g a t a c g t t c t t t g g a t t a t g a t t t t a a g a a a t t a t t a a t t t g a a
a t g a t a g g t a a g g a a g c t g g a t t a t c t c c t g t t g a g a c a a g a g a t t c g t t t t g a c a t a g a t g a a t g
c a t t t t c c c c t c t c c t c c c t g c t a c c a t t a t a t t t t g g g t t a t g t t t t g c t t c t t a a g a t a g a a a
t c c c a g t t c t c t a a t t t g g t t t t c t t t t g g g a a c c a a c a t a c a a a t g a a t c a g t a t c a a t t a g g g c
c t g g g g t a g a g a c a g a a a c t t g a g a g a a g a a g t t a g t g a t t c c c t c t c t t t c t a g t t t g g t a g g a a
t c a c c c t g a a g a c c t a g t c c t c a a t t t a a t t g t g t g g g t t t t t a a t t t t c c t a g a a t g a a g t g a c t g a a a
c a a t g a g a a g a a t a c a g c a c a a c c c t t g a a c a a a t g t a t t t a g a a a t a t a t t t a g t t t t a t a g c a g a a
g c a g c t c a a t t g t t t g g t t g g a a a g t a g g g g a a a t t g a a g t t g t a g t c a c t g t c t g a g a a t g g c t a t g a a
g c g t c a t t t c a c a t t t t a c c c a a c t g a c c t g c a t g c c c a g g a c a c a a g t a a a a c a t t t g t g a g a t a g t g
g t g g t a a g t g a t g c a c t c g t g t t a a g t c a a a g g c t a t a a g a a c a c t g t g a a a g t t c a t a t t c a t c c a t
t g t g a t t c t t t c c c a c g t c t t g c a t g t a t t a c t g g a t t c c c a c a g t a a t a g a c t g t g c a t g g t g t g t
a t a t t t c a t t g c g a t t t c c t g t t a a g a t g a g t t t g t a c t c a g a a t t g a c c a a t t c a g g a g g t g t a a a a a t
a a a c a g t g t t c t c t c t a c c c c a a a g c c a c t a c t g a c c a a g g t c t c t t c a g t g c a c t c g c t c c c t c t c
t g g c t a a g g c a t g c a t t a g c c a c t a c a c a a g t c a t t a g t g a a a g t g g t c t t t a t g t c c t c c c a g c a g a c
a g a c a t c a a g g a t g a g t t a a c c a g g a g a c t a c t c c t g t g a c t g t g g a g c t c t g g a a g g c t t g g t g g g a g t
g a a t t t g c c c a c a c c t t a c a a t t g t g g c a g g a t c c a g a a g a g c c t g t c t t t t a t a t c c a t t c c t t g a t g
t c a t t g g c c t c t c c c a c c g a t t t c a t t a c g g t g c c a c g c a g t c a t g g a t c t g g g t a g t c o g g a a a c a a a
a g g a g g g a a g a c a g c c t g g t a a t g a a t a a g a t c c t t a c c a c a g t t t t c t c a t g g g a a a t a c a t a a t a a a c
c c t t t c a t c t t t t t t t t t t c c t t t a a g a a t t a a a a c t g g g a a t a g a a c a t g a a c t g a a a a g t c t t g c
a a t g a c a a g a g g t t t c a t g g t c t t a a a a g a t a c t t t a t a t g g t t g a a g a t g a a t c a t t c c t a a a t t a a
c c t t t t t t t t a a a a a a a a c a a t g t a t a t t a t g t t c c t g t g t g t t g a a t t t a a a a a a a a a a a a a c t t t a
c t t g g a t a t t c a t g t a a t a t a a a g g t t t g g t g a a a t g a a c t t t a g t t a g g a a a a g c t g g c a t c a g c t
t t c a t c t g t g t a a g t t g a c a c c a a t g t g t c a t a a t a t t c t t t a t t t t g g g a a a t t a g t g t a t t t t a t a a a
a a t t t t a a a a g a a a a a a g a c t a c a c a g g t t a a g a t a a t t t t t t a c c t g t c t t t t c c a t a t t t t a a
g c t a t g t g a t t g a a g t a c c t c t g t t c a t a g t t t c c t g g t a t a a a g t t g g t t a a a a t t t c a t c t g t t a a t a
g a t c a t t a g g t a a t a t a a t g t a t g g g t t t t c a t t g g t t t t t g c a g a c a g t a g a g g g a g a t t t t g t a a c
a a g g g c t t g t t a c a c a g t g a t a t g g t a a t g a a a a t t g c a a t t a t c a c t c c t t t t c a t g t t a a t a a t t

tgaggactggataaaaaggtttcaagattaaaatttgatgttcaaaccctttgt

Figure 43

Name: Similar to
ID No. 44

Gen Bank Accession No. AK021429

SEQ

MUS Musculus
Plenty of SH3S (PRT3)

```

1  ctgagagaca ctgagagcgg cgagcgcggg ggggcccgcg ctgcatcagc cgccgcagcc
61  gctgcggggc cgcgaaacaa gaggaggagc cgaggcgcga gagcaaagtc tgaaatggat
121 gttacatgag tcattttaag ggatgcacac aactatgaac atttctgaag attttttctc
181 agtaaagtag ataaagatgg atgaatcagc cttgttggat cttttggagt gtccggtgtg
241 tctagagcgc cttgatgctt ctgcgaaggt cttgccttgc cagcatcagc tttgcaagcg
301 atgtttgctg gggatcgtag gttctcgaaa tgaactcaga tgtcccgagt gcaggactct
361 tgttggctcg ggtgtcgagg agcttcccag taacatcttg ctggtcagac ttctggatgg
421 catcaaacag aggccttgga aacctgggtc tggtggggga agtgggacca actgcacaaa
481 tgcattaagg tctcagagca gcactgtggc taattgtagc tcaaaagatc tgcagagctc
541 ccagggcgga cagcagcctc ggggtgcaatc ctggagcccc ccagtgaggg gtatacctca
601 gttaccatgt gccaaagcgt tatacaacta tgaaggaaaa gagcctggag accttaaatt
661 cagcaaaggc gacatcatca ttttgccaag acaagtggat gaaaattggg accatgggga
721 agtcaatgga atccatggct ttttcccac caactttgtg cagattatta aaccgttacc
781 tcagccccc cctcagtgca aagcacttta tgactttgaa gtgaaagaca aggaagcaga
841 caaagattgc cttccatttg caaaggatga tgttctgact gtgatccgaa gagtggatga
901 aaactgggct gaaggaatgc tggcagacaa aataggaata ttccaattt catatgttga
961 gtttaactcg gctgctaagc agctgataga atgggataag cctcctgtgc caggagttga
1021 tgctggagaa tgttcctcgg cagcagccca gagcagcact gccccaaagc actccgacac
1081 caagaagaac accaaaaagc ggcactcctt cacttccctc actatggcca acaagtctc
1141 ccaggcatcc cagaaccgcc actccatgga gatcagcccc cctgtcctca tcagctccag
1201 caaccccact gctgctgcac ggatcagcga gctgtctggg ctctcctgca gtgccccttc
1261 tcagggttcat ataagtacca ccgggttaat tgtgaccccc ccccaagca gccagtgac
1321 aactggcccc tcgtttactt tcccatcaga tgttccctac caagctgccc ttggaacttt
1381 gaatcctcct cttccaccac cccctctcct ggctgccact gtccttgctt ccacaccacc
1441 aggcgccacc gccgcgctg ctgctgctgg aatgggaccg aggcccatgg caggatccac
1501 tgaccagatt gcacatttac ggccgcagac tcgccccagt gtgtatgttg ctatatatcc
1561 atacactcct cggaaagagg atgaactaga gctgagaaaa ggggagatgt ttttagtggt
1621 tgagcgctgc caggatggct ggttcaaagg gacatccatg cataccagca agataggggt
1681 tttccctggc aattatgtgg caccagtcac aagggcggtg acaaatgctt cccaagctaa
1741 agtccctatg tctacagctg gccagacaag tcggggagtg accatggtca gtccttccac
1801 ggcaggaggg cctgccaga agctccaggg aaatggcgtg gctgggagt cagtggtgtg
1861 cccgcagct gtggtatcag cagctcacat ccagacaagt cctcaggcta aggtcttgtt
1921 gcacatgacg gggcaaatga cagtcaacca ggcccgaat gctgtgagga cagttgcagc
1981 gcacaaccag gaacgcccc cggcagcagt gacacccatc caggtacaga atgccgccgg
2041 cctcagccct gcactgtgtg gcctgtccca tcaactcgtg gcctcccac aacctgcgcc
2101 tctgatgcca ggctcagcca cgcacactgc tgccatcagt atcagtcgag ccagtgcgcc
2161 tctggcctgt gcagcagctg ctccactgac ttccccaagc atcaccagt cttctctgga
2221 ggctgagccc agtggccgga tagtgaccgt tctccctgga ctcccacat ctctgacag
2281 tgcttcatca gcttgtggga acagttcagc aaccaaacca gacaaggata gc

```

FIGURE 43

Figure 44

Name: SH3-Gen Bank Accession No. U26710SEQID No. 45

Binding Protein

CBL-B

```

1  ctgggtcctg tgtgtgccac aggggtgggg tgtccagcga gcggtctcct cctcctgcta
61  gtgctgctgc ggcgtcccgc ggcctcccgc agtcgggcgg gaggggagag cgggtgtgga
121 tttgtcttga cggtaattgt tgcgtttcca cgtctcggag gcctgcgcgc tgggttgctc
181 cttcttcggg agcgagctgt tctcagcgat ccactccca gccggggctc cccacacaca
241 ctgggctgcg tgcgtgtgga gtgggaccgc cgcacacgcg tgtctctgga cagctacggc
301 gccgaaagaa ctaaaattcc agatggcaaa ctcaatgaat ggcagaaacc ctggtggtcg
361 aggaggaaat ccccgaaaag gtcgaatttt ggggtattatt gatgctattc aggatgcagt
421 tggacccccct aagcaagctg ccgcagatcg caggaccgtg gagaagactt ggaagctcat
481 ggacaaaagt gtaagactgt gccaaaatcc caaacttcag ttgaaaaata gccaccata
541 tatacttgat attttgctg atacatatca gcatttacga cttatattga gtaaatatga
601 tgacaaccag aaacttgccc aactcagtga gaatgagtac tttaaaatct acattgatag
661 ccttatgaaa aagtcaaaac gggcaataag actctttaaa gaaggcaagg agagaatgta
721 tgaagaacag tcacaggaca gacgaaatct cacaaaactg tcccttatct tcagtcacat
781 gctggcagaa atcaaagcaa tctttcccaa tgggtcaattc cagggagata actttcgtat
841 cacaaaagca gatgctgctg aattctggag aaagtttttt ggagacaaaa ctatcgtacc
901 atggaaaagta ttcagacagt gccttcatga ggtccaccag attagctcta gcctggaagc
961 aatggctcta aaatcaacaa ttgatttaac ttgcaatgat tacatttcag ttttgaatt
1021 tgatattttt accaggtgtt ttcagccttg gggctctatt ttgcggaatt ggaatttctt
1081 agctgtgaca catccaggtt acatggcatt tctcacatat gatgaagtta aagcagcagt
1141 acagaaatat agcaccaaac ccggaagcta tattttccgg ttaagttgca ctcgattggg
1201 acagtgggcc attggctatg tgactgggga tgggaatatc ttacagacca tacctcataa
1261 caagccctta tttcaagccc tgattgatgg cagcagggaa ggattttatc tttatcctga
1321 tgggaggagt tataatcctg atttaactgg attatgtgaa cctacacctc atgaccatat
1381 aaaagttaca caggaacaat atgaattata ttgtgaaatg ggctccactt ttcagctctg
1441 taagatttgt gcagagaatg acaaagatgt caagattgag ccttggtggc atttgatgtg
1501 cacctcttgc cttacggcat ggcaggagtc ggatggtcag ggctgccctt totgtcgttg
1561 tgaaataaaa ggaactgagc ccataatcgt ggaccccttt gatccaagag atgaaggctc
1621 caggtgttgc agcatcattg accccttttg catgccgatg ctgactttgg acgacgatga
1681 tgatcgtgag gagtccctga tgatgaatcg gttggcaaac gtccgaaagt gcactgacag
1741 gcagaactca ccagtcacat caccagatca ctctcccctt gccagagaa gaaagccaca
1801 gcctgaccca ctccagatcc cacatctaag cctgccaccc gtgcctcctc gcctggatct
1861 aattcagaaa ggcatagtta gatctccctg tggcagccca acaggttcac caaagcttc
1921 tccttgcatg gtgagaaaaa aagataaacc actcccagca ccacctcctc ccttaagaga
1981 tcctcctcca ccgccacctg aaagacctcc accaatccca ccagacaata gactgagtga
2041 acacatccat catgtggaaa gcgtgccttc cagagacccg ccaatgcctc ttgaagcatg
2101 gtgccctcgg gatgtgtttg ggactaatca gcttgtggga tgtcgactcc taggggaggg
2161 ctctccaaaa cctggaatca cagcgagttc aaatgtcaat ggaaggcaca gtagagtggg
2221 ctctgaccca gtgcttatgc ggaaacacag acgccatgat ttgcctttag aaggagctaa
2281 ggtcttttcc aatggtcacc ttggaagtga agaatatgat gttcctcccc ggctttctcc
2341 tcctcctcca gttaccaccc tcctccctag cataaagtgt actggtccgt tagcaaattc
2401 tctttcagag aaaacaagag acccagtaga ggaagatgat gatgaatata agattccttc
2461 atcccaccct gtttccttga attcacaacc atctcattgt cataatgtaa aacctcctgt
2521 atgggtcgtg gataatggtc actgtatgct gaatggaaca catggtccat cttcagagaa
2581 gaaatcaaac atccctgact taagcatata tttaaaggga gatgtttttg attcagcctc
2641 tgatcccgty ccattaccac ctgccaggcc tccaaactcg gacaatccaa agcatggttc
2701 ttcactcaac aggacgccct ctgattatga tcttctcatc cctccattag gtgaagatgc
2761 ttttgatgcc ctccctccat ctctcccacc tccccacct cctgcaaggc atagtctcat
2821 tgaacattca aaacctcctg gctccagtag ccggccatcc tcaggacagg atctttttct
2881 tcttccttca gatccctttg ttgatctagc aagtggccaa gttcctttgc ctctgctag
2941 aaggttacca ggtgaaaatg tcaaaaacta cagaacatca caggactatg atcagcttcc

```

```
3001 ttcatgttca gatggttcac aggcaccagc cagaccccct aaaccacgac cgcgcaggac
3061 tgcaccagaa attcaccaca gaaaacccca tgggcctgag gcggcattgg aaaatgtcga
3121 tgcaaaaatt gcaaaactca tgggagaggg ttatgccttt gaagaggtga agagagcctt
3181 agagatagcc cagaataatg tcgaagttgc ccggagcatc ctccgagaat ttgccttccc
3241 tcctccagta tccccacgtc taaatctata gcagccagaa ctgtagacac caaaatggaa
3301 agcaatcgat gtattccaag agtgtggaaa taaagagAAC tgagatggaa ttcaagagag
3361 aagtgtctcc tcctcgtgta gcagcttgag aagaggcttg ggagtgcagc ttctcaaagg
3421 agaccgatgc ttgctcagga tgtcgacagc tgtggcttcc ttgtttttgc tagccatatt
3481 tttaaatcag ggttgaactg acaaaaataa tttaaagacg tttacttccc ttgaactttg
3541 aacctgtgaa atgctttacc ttgtttacaa tttggcaaag ttgcagtttg ttcttgtttt
3601 tagtttagtt ttgttttggt gttttgatac ctgtactgtg ttcttcacag accctttgta
3661 gcgtgggtcag gtctgctgta acatttccca ccaactctct tgctgtccac atcaacagct
3721 aaatcattta ttcatatgga tctctaccat ccccatgcct tgcccaggtc cagttccatt
3781 tctctcattc acaagatgct ttgaaggttc tgattttcaa ctgatcaaac taatgcaaaa
3841 aaaaaaagta tgtattcttc actactgagt ttcttctttg gaaaccatca ctattgagag
3901 atgggaaaaa cctgaatgta taaagcattt atttgtcaat aaactgcctt ttgtaagggg
3961 ttttcacata aaaaaaaaaa aa
```

FIGURE 44

Figure 45

Name: SH3-
46

Gen Bank Accession No. AB028645

SEQ ID No.

Binding Protein
CBL-C

```

1  tcccagccgc accggtcctt cccggcacac gcgaggctcc catggctctg gcgggtggccc
61  cgtgggggcg acagtgggaa gaggcccgcg ccttggggcg ggcagtcagg atgtctgcagc
121  gcctagaaga gcaatgcgtc gacccccggc tgtccgtgag tcccccttcg ctgcgggacc
181  tgctgccccg cacagcgcag ctgcttcgag aggtggccca ttctcggcgg gcggccggcg
241  gaggcgggccc cgggggtccc ggcggtctct gggactttct actcatctac ctggccaatc
301  tggaggccaa gagcaggcag gtggccgcgc tgctgcctcc ccggggccga aggagtgcc
361  acgacgagct cttccggcg ggctccagac tcaggcgaca gctggccaag ctggccatca
421  tcttcagcca catgcacgca gagctgcacg cactcttccc cgggggaaag tactgtggac
481  acatgtacca gctcaccaag gcccccgccc acaccttctg gagggaaaag tgcggagccc
541  ggtgtgtgct gccctgggct gagtttgagt ccctcctggg cacctgccac cctgtggaac
601  caggctgcac agccctggcc ttgcgcacca ccattgacct cacctgcagc gggcacgtgt
661  ccattctcga gttcgacgtc ttcaccaggc tctttcagcc atggccaaca ctccctcaaga
721  actggcagct cctggcagtc acccaccag gctacatggc ctccctcacc tatgatgagg
781  tccaagagcg tctgcaggcc tgcagagaca agccaggcag ttacatcttc cggcccagct
841  gtactcgctt ggggcagtg gcatcggt atgtgagctc agatggcagc atcctgcaga
901  ccattccctg caacaaaccc ctgtcccagg tgctcctgga gggacagaag gacggcttct
961  acctctaccc agatggaaag acccacaacc cagacctgac tgagctcggc caggcagaac
1021  cccagcagcg catccacgtg tcagaggagc agctgcagct ctactgggccc atggactcca
1081  catttgagct ctgcaagatc tgtgctgaga gcaacaagga tgtgaagatt gagccgtgcg
1141  ggcacctgct ctgcagctgc tgcttggtg cctggcagca ctcgacagc cagacctgcc
1201  ccttctgccg ctgcgagatc aagggtggg aggcctgag tatctaccag ttctacggtc
1261  aggctactgc tgaggactca gggaacagca gtgaccagga aggcaggag ttggagctgg
1321  ggcaggtgcc ctttcggct cctccattgc cccacggcc agatctgccc cccaggaagc
1381  ccagaaatgc ccagccgaaa gtgagactcc taaaggggaa ctcccctcca gctgcgctgg
1441  gacccagga ccctgccccg gcctgaaggc cagggcacc agatgtgctg ctcaaggag
1501  cccaagggc tgaaggggg ttgtgaaacc gaaataaact gccaagcctg gtctgtcaaa
1561  aaaaaaaaaa aaaaa

```

FIGURE 45

Figure 46

Name: MultipleGen Bank Accession No. AF064801SEQID No. 47Membrane Spanning
Receptor TRC8

```

1  gaaaccacaga gacctcctgg ggagccgccg ccgcgcgcct ctgggccatc gctgcctccg
61  ccgcctgctc caccctcgagg gacgcgagcg ggccggcgggg ctggcccgta gagagacagg
121 agaggaagga gggcaggggc ggagttgccc gccttagccc ccgcccccg ccgcggcccc
181 gggccctgcc ccgcgcggcc ctgcccggcc caccgagccc tgggtgtggca gcggctcatg
241 gcggccgctg ggcccccgca gcagcaggtg cggatggccc atcagcagat ctgggcggcg
301 ctogaagtgg cgctccgggt gccctgcctt tacatcatcg acgccatctt caactcctac
361 ccggattcca gccaaagccg gttctgcata gtgctccaga tcttctcccg gctctttggt
421 gtatttgcat ccagtattgt tctgatcttg tcacaacgat cacttttcaa gttttacacg
481 tacagctcag ctttctgtt agctgcaact tcagtgttg tgaattatta tgcttctttg
541 cacattgact tctatggtgc ctacaacacg tcagcttttg gaattgagct gcttctcga
601 aaaggctcct cgctgtggat ggcacttata gttctacagc taacatttgg aattggatac
661 gttacactac tccagattca ttccatctat tcacaattaa ttattttgga tctcttggtt
721 cctgtaatag gcttaatcac agagctacca ttacacatca gagagacttt actgtttact
781 tcttcttga ttctcacatt aaatacagtg tttgtcctgg cagtgaact gaagtgggtt
841 tattattcca cagatatgt ttatcttttg gtgaggcaca tgtatcgaat ttatggatta
901 cagttattga tggaggacac atggaagagg attcgtttcc cagacatact acgagctctt
961 tggctaacaa gagttacagc tcaggctaca gtgttaatgt acatcttaag gatggcaaat
1021 gaaactgatt cttcttttat ttcttgggat gatttttggg acctcatttg caatcttata
1081 attagtgggt gcgattctac actaactgta ctgggcatga gtgctgtaat ttctcagta
1141 gccattatt tggggcttgg aatattggcc tttattggat caactgagga agatgacagg
1201 cgtcttggct ttgttgacac tgttttattt tttatttttg ctcttcagac tgggttaagt
1261 gggctaagac cagaagagag acttattcgc ttaagtagaa acatgtgcct tttattaact
1321 gcagtcctgc attttatcca tggaaatgaca gaccctgtat taatgtctct cagtgcctct
1381 catgtgtcat ctttctgtag acattttcct gtgctgtttg tctctgcttg cctgtttatt
1441 cttcctgtct tactcagtta tgttcttttg catcactatg cactaaatac atgggtgttt
1501 gcagttacag catttttgtgt ggaactgtgc ttaaaagtaa ttgtttctct cactgtttat
1561 acgttattca tgattgatgg ctactataat gtccctctggg aaaagcttga cgattatgtc
1621 tactacgttc gttcaacagg cagtattatt gaatttatat ttggagttgt aatgtttgga
1681 aatggggctt acactatgat gtttgagtcg ggaagtaaaa ttccggcttt tatgatgtgc
1741 ctacatgcat attttaacat ctacttacia gccaaaaatg gctggaagac atttatgaat
1801 cgtaggactg ctgtgaagaa aattaattca cttcctgaaa taaaaggag ccgcttacia
1861 gaaataaatg atgtatgtgc aatctgctat catgagtta caacatctgc tcgtattaca
1921 ccgtgtaatc attatttcca tgcactttgc cttcggaaat ggctgtacat tcaagatact
1981 tgccaatgt gccatcagaa agtatacatc gaagatgata tcaaggataa ttcaaatgta
2041 tctaacaaca atggatttat tccaccaat gaaactccag aggaagctgt aagagaagct
2101 gctgctgaat ctgacaggga attgaacgaa gatgacagta cagatttgta tgatgatgtt
2161 caaagagaaa gaaatggagt gattcagcac acaggcgcag cagctgaaga atttaatgat
2221 gatactgact gatgaaaata gcatttatta atgattgagg tatttgttta aaattcagtt
2281 catccaaaat ggagtaatat cttcacctt cagtgtgtaa ccaagcacia aaacagtatc
2341 aatgttgaat ctgtgaatgg ttttccgttt actgtgatgt gctactgtaa atatacctct
2401 ttaattactt ctggtctctt tgggtgacct tttaaatttg tgtacattat tgtacataga
2461 ataaaatggt ttcacatttt tatgacaaaa aaaaaaaaaa aaaaa

```

FIGURE 46

Figure 47

Name: Autocrine	Gen Bank Accession No. AF124145				SEQ ID No.	
48						
Motility Factor Receptor						
1	gggccgcgcg	agaggcccg	ccgcagcgca	gggaagcctg	ggggccagag	gtcgccgctg
61	ccgccatgcc	gctgctcttc	ctcgagcgct	tcccctggcc	cagcctccgc	acctacacgg
121	gcctcagcgg	cctggccctg	ctgggcacca	tcatcagcgc	ctaccgcgcg	ctcagccagc
181	ccgagggcgg	ccccggcgag	ccggaccagc	taacggcctc	gctgcagcct	gagccgccgg
241	cgcccgcccc	gccgagcgcc	gggggacccc	ggggccgcga	tgtggcccag	tacctgctct
301	cagacagcct	cttcgtgtgg	gttctagtaa	ataccgcttg	ctgtgttttg	atgttggtgg
361	ctaagctcat	ccagtgtatt	gtgtttggcc	ctcttcgagt	gagtgcagag	cagcatctca
421	aagacaaaatt	ttggaatttt	atcttctaca	agttcatttt	catctttggt	gtgctgaatg
481	tccagacagt	ggaagaggtg	gtcatgtggt	gcctctgggt	tgccggactt	gtctttctgc
541	acctgatgg	tcagctctgc	aaggatcgat	ttgaatatct	ttccttctcg	cccaccaacg
601	cgatgagcag	ccacggctga	gtcctgtccc	tgttggttgc	catgctgctt	tctgctgttg
661	gactggcggc	cgtctgctcc	atcaccggct	acaccacgg	aatgcacacc	ttggctttca
721	tggctgcaga	gtctcttctt	gtgacagtga	ggactgctca	tgtgatttta	cgatacgtaa
781	ttcacctctg	ggacctcaac	cacgaaggga	cgtgggaagg	aaaggggacg	tatgtctatt
841	acacagactt	tgtcatggag	ctcactctcc	tgtccctgga	cctcatgcac	catattcaca
901	tgttggttatt	tggcaacatc	tggttatcca	tggccagcct	ggtcatcttt	atgcagctgc
961	gttacctgtt	tcatgaggtg	caacgtcgaa	ttcgtcggca	caagaactat	ctacgtgtgg
1021	ttggaacat	ggaggccagg	tttgcagttg	caactccaga	ggagctggct	gtcaacaatg
1081	acgactgtgc	catctgttgg	gactccatgc	aggctgcgcg	gaaactgccc	tgtggacatc
1141	ttttccacaa	ctcctgtctt	cgttcctggc	tagaacaaga	cacctcctgt	ccaacatgca
1201	gaatgtctct	taatattgcc	gacaataatc	gtgtcagggg	agaacatcaa	ggagagaact
1261	tggatgagaa	tttggttcct	gtagcagcag	ccgaagggag	acctcgctta	aaccaacaca
1321	atcacttctt	ccatttcgat	gggtctcgga	ttgcgagctg	gctgccgagt	ttttcggttg
1381	aagtgatgca	caccaccaac	attcttggca	ttacgcaggc	cagcaactcc	cagctcaatg
1441	caatggctca	tcagattcaa	gagatgtttc	cccaggttcc	ataccatctg	gtactgcagg
1501	acctccagct	gacacgctca	gttgaaataa	caacagacaa	tattttagaa	ggacggattc
1561	aagtaccttt	tcctacacag	cggtcagata	gcacagacc	tgcattgaac	agtcctgtgg
1621	aaaggccaag	cagtgaccag	gaagagggag	aaacttctgc	tcagaccgag	cgtgtgccac
1681	tggacctcag	tcctcgccctg	gaggagacgc	tggacttcgg	cgaggtggaa	gtggagccca
1741	gtgaggtgga	agacttcgag	gctcgtggga	gccgcttctc	caagtctgct	gatgagagac
1801	agcgcagct	ggtgcagcgt	aaggacgaac	tcctccagca	agctcgcaaa	cgtttcttga
1861	acaaaagtct	tgaagatgat	gcggcctcag	agagcttctc	ccccttggaa	ggtgcgtcct
1921	ctgaccccg	gacctgcgt	cgaaggatgc	tggctgccgc	cgcggaacgg	aggcttcaga
1981	agcagcagac	ctcctagcgc	tcccttgctc	tcctcagctg	cctcctgcgc	cctgtgcccg
2041	actgactgga	ggaggcctgt	cccaattctg	cccgtccat	ggaaaagcgg	gcttgactgc
2101	attgccgctg	tataaagcat	gtggtcttat	agtgtttgga	cagctgataa	atttaatcct
2161	tcctttgta	actttctatg	tgacatttct	cttccccctta	gaaacactgc	aaattttaac
2221	tgtaggtatg	atctcttctg	gtgttgactg	gactgcttgg	ggtggggggc	gatcaggagg
2281	aagtgagcag	tcgctgcct	gcagcaggca	gctttctact	cctgcctcat	gcatacgtcc
2341	cacaaatgca	ggtgtcctga	gcaccacacc	cagtgggaag	agtgtggggg	aggcgcacag
2401	tgtgagcccg	ccccacgctc	gtggggtaac	atctgttatc	aaactgctgt	cgttggttgtg
2461	gaagcatgta	gactgtgcc	gaggccagac	ccacgggctc	atgcaccctt	gagccagcag
2521	ggcatcttgg	aaaaggaact	cttggttcga	tacctggagc	agaggagggg	aaagtccagg
2581	gctataggg	gtgatgaagt	cacccttttc	tgtcccacta	catctggggac	tgactttccg
2641	agcctccagt	ccaaagccgg	cttgattttc	gtgaactctg	gtgctcctgc	atctcatgag
2701	tgtgccccat	gggtcccctc	ccctctcagc	atttccctgt	cccgtctgga	cctggggggg
2761	tggtaggca	gcaagctttg	gtttatggtt	ttcattcatt	ggtgaagtaa	attaggcagt
2821	gctaaagcct	gtgggtttgg	tccttgaaca	agatgtgggc	cttgcaagat	gggagagtaa
2881	accttgaagg	gctttatt				

FIGURE 47

Figure 48

Name: MembraneGen Bank Accession No. AF255303SEQ ID No.49

Associated Nucleic Acid

Binding Protein (Fragment)

1 agccaagtct tgtcagagat ttctcttttc aggtggcaaa gctgttttct tcacacttga
61 gtctctacaa tattgttttg atcagtagtt tocaaagtcc attaacctcct ggccatactt
121 tattatgttt tggggacttg gttatccaag ggaaacactt ttttaacaa caaaacaaaa
181 aaaccgcca gcagtcacaa gtaatttgtg ttcttaaaaa tggaatatgg aaagttaatt
241 tgcttgtttg atgtggctgt tgagaaaaat acataaaagc tttgatgttt attatgtgag
301 caaccaatat aaatacagtt tagttgaaag gaacactatt aaggtattgt ttccaggcag
361 aatttcagaa atgtaattaa ttcagcaaat aggtttttta aaaaagacat ccaaaggtta
421 taaaattatt tagaagtatt ttaggtctga agctgtaata gttgacttaa gcaattaact
481 cttcaaaggt gaatgatgaa tatgtgggta attcatactt ttgtccattt ctagcttaca
541 aaacactaca cagcaaaata atgatctgct agactgctaa cccgagcatc cagcttcac
601 aatgcctgtg caggcagctc aatggacaga atttctgtcc tgtccaatct gctataatga
661 atttgatgag aatgtgcaca aacccatcag tttagggtgt tcacacactg tttgcaagac
721 ctgcttgaat aaacttcac gaaaagcttg tccttttgac cagactgcc acaacacaga
781 tattgatgta ctctctgtca acttcgact tctccagtta gttggagccc aggtaccaga
841 tcacagtcata attaatgttaa gtaatctagg tgagaataaa cactatgagg ttgcaaagaa
901 atgcgttgag gatttggcac tctacttaaa accactaagt ggaggtaaag gtgtagctag
961 cttgaaccag agtgcaactga gccgtccaat gcaaaggaaa ctggtgacac ttgtaaactg
1021 tcaactggtg gaggaagaag gtcgtgtaag agccatgcga gcagctcgtt cccttggaga
1081 aagaactgta acagaactga tattacagca ccagaacct cagcagttgt ctgccaatct
1141 atgggcccgt gtcaggcctc gaggatgcca gtttttaggg ccagctatgc aagaagaggc
1201 cttgaagctg gtgttactgg cattagaaga tggttctgcc ctctcaagga aagttctggt
1261 actttttgtt gtgcagagac tagaaccaga atttcctcag gcatcaaaaa caagtattgg
1321 tcatgttgtg caactactgt atcgagcttc ttgttttaag gttaccaaaa gagatgaaga
1381 ctcttcctta atgcagctga aggaggaatt tcggagttat gaagcattac gcagagaaca
1441 tgatgcccaa attgttcata ttgccatgga agcaggactc cgtatttcac ctgaacagt
1501 gtctctctct ttgtatgggtg atttggctca taaatcacac atgcagtcta tcattgataa
1561 gctacagtct ccagagtcac ttgcaaagag tgtccaggaa ttgacaattg ttttgcaacg
1621 aacaggtgac ccagctaact taaatagact gaggcctcat ttagagcttc ttgcaaacat
1681 agaccctaatt ccagacgctg tttcaccaac ttgggagcag ctggaaaatg caatggtagc
1741 tgttaaaaca gtagttcatg gccttgtgga cttcatacaa aattatagta gaaaaggcca
1801 tgagaccctc cagcctcagc caaacagcaa atacaagact agcatgtgcc gagatttgcg
1861 acagcagggg ggttgtccac gaggaacaaa ttgtacattt gccattctc aggaagagct
1921 tgaaaagtat cgattaagga acaaaaagat caatgccact gtaagaacgt ttctctctct
1981 aaataaagtt ggtgtaaaca acactgtcac aaccacagcc ggaaatgtca tttctgtcat
2041 aggaagtact gaaacaacag ggaaaattgt tccaagtaca aacggaattt caaatgcaga
2101 aaacagtgtt tccagctaa tctcacgtag tactgacagt accttaagag ccttgagac
2161 cgtgaagaaa gtgggaaagg ttggcgctaa tggtcagaat gctgctgggc cctctgcaga
2221 ttctgtaact gaaaataaaa ttggttctcc acccaagact cctgtaagta atgtagcagc
2281 tacctcagct gggccctcta atgttggaac agagctgaat tctgtgcctc aaaaatccag
2341 cccatttcta actagagtac cagtatatcc tccgcattct gaaaacattc agtattttca
2401 agatccaagg actcagatac cctttgaagt cccacagtac ccacagacag gatactatcc
2461 accacctcca acggtaccag ctggtgtggc tccctgtgtt cctcgctttg tgaggtccaa
2521 taacgttcca gagtcctccc tcccactgc ttccatgcca tatgccgatc attacagtac
2581 attttcccct cgagatcgaa tgaattcttc tccttaccag cctcctcctc cgcagccgta
2641 tggaccagtt cctccagtac cttctggaat gtatgctcct gtgtacgaca gcaggcgcat
2701 ctggcgccca cctatgtacc aacgagatga cattattaga agcaattctt tacctccaat
2761 ggatgtgatg cactcatctg tctatcagac atctttgcgg gaaagatata actcattaga
2821 tggatattat tcggtggctt gtcagccacc aagtgagcca aggacaactg tgccctttacc
2881 aagggaacct tgtgttcatt tgaagaccag ttgcgaggag cagataagaa gaaagccaga
2941 tcagtgggca cagtaccaca ctcagaaagc acctcttgct tcttcaactc ttctgtgagg

```
3001 aacacagtca ccaacaccac cttctcctct gttcagtgtg gactttcgtg cggatttctc
3061 agagagtgtg agtgggtacaa aatttgaaga agatcatctt tccattatt ctccctgggc
3121 ttgtggcacc ataggctcct gtataaatgc cattgattca gagcccaaag atgtcattgc
3181 taattcaaat gctgtgttaa tggacctgga cagtgggtgat gttaagagaa gagtacattt
3241 atttgaaacc cagagaagga caaaagaaga agatccaata attcccttta gtgatggacc
3301 catcatctca aaatgggggtg cgatttccag atcttcccgt acagggtacc ataccacaga
3361 tcctgtccag gccactgctt cccaaggaag tgcgactaag cccatcagtg tatcagatta
3421 tgtcccttat gtcaatgctg ttgattcaag gtggagttca tatggcaacg aggccacatc
3481 atcagcacac tatgttgaaa gggacagatt cattgttact gatttatctg gtcataaaaa
3541 gcattccagt actggggacc ttttgagcct tgaacttcag caggccaaga gcaactcatt
3601 acttcttcag agagaggcca atgctttggc catgcaacag aagtggaatt ccctggatga
3661 aggccgtcac cttaccttaa accttttaag caaggaaatt gaactaagaa atggagagtt
3721 acagagtgat tatacagaag atgca
```

FIGURE 48

Figure 49

Name: NADH-
ID No. 50

Gen Bank Accession No. X81900

SEQ

Ubiquinone Oxidoreductase
MWFE Subunit

```
1 taggtaacgg ggcagagatg tggttcgaga ttctccccgg actctccgtc atgggcgtgt
61 gcttggttgat tccaggactg gctactgctg acatccacag gttcactaac gggggcaagg
121 aaaaaagggg tgctcatttt gggatcactt ggagtctgat ggaaagagat aggcgcacat
181 ctggagttga tcgttactat gtgtcaaagg gtttggagaa cattgattaa ggaagcattt
241 tcctgattga tgaaaaaat aactcagtta tggccatcta ccctgctag aaggttacag
301 tgtattatgt agcatgcaat gtgttatgta gtgcttaata aaaataaaat ga
```

FIGURE 49

Figure 50

Name: Peroxisome Gen Bank Accession No. M86852 SEQ ID No.
51
Assembly Factor-1
(PAF-1)

```
1 ccatccagtgc ccttccgcag cgccgctaaa ggcgagttct cgttgggtgta actttttctt
61 ttttttttca gccacttccg gctcctgcgt cgctccggaa gcctgcgagt tccggaagcc
121 ttggtaatcc agattcggct aggaaaagac aagctttcca gagaatgttt cagagaaagt
181 tacgtggagc gtgggcgttt cgcagactcc taagtggctt ggaacctaac ctgcagtgtc
241 tctgagcttc tggcaagtat gaggcaagga tttcacagaa gaacttggag caaaaattcc
301 cactgtcata gctataggct gtccagtcct gagggagctg caggaacagg aaaaagagac
361 cttcagagaa gacatggctt ccagaaaaga gaatgcgaag agtgcaaaca gactgctaag
421 aataagccag ttggatgcac ttgaactaaa caaggccctg gagcagctag tttgggtcca
481 gtttactcag tgctttcatg gatttaaacc tgggctgtta gctcgctttg agccagaggt
541 gaaagcgtgc ttatgggttt tcttgtggag attcaccatc tactccaaaa atgccacagt
601 gggacagtca gttttgaata ttaagtacaa aaatgatttt tcccctaacc tgagatatca
661 gccaccagcgt aaaaatcaaa aaatctggta tgctgtttgt acaattgggtg gcaggtgggt
721 agaagaacga tgctatgatt tgtttcgaaa ccatcattta gcatcatttg ggaaagtcaa
781 gcagtgtgtg aattttgtga ttggactttt gaaattaggt gggctgatta attttttgat
841 tttccttcag aggggaaagt ttgcaacttt gacagaacgt ctctaggtta ttcattctgt
901 attttgcaag cctcaaaaaca tacgtgaagt tggctttgaa tacatgaata gggaaacttct
961 ctggcatggg tttgctgaat ttctgatttt tctcttacca cttatcaatg tccagaagtt
1021 gaaagccaag ctgtcttcat ggtgtattcc tcttactggg gcacctaata gtgacaatac
1081 attagccacc agtggcaaag aatgcgctct atgtggagag tggcccacca tgcctcacac
1141 cataggatgt gagcatattt tctgttattt ctgtgctaag agtagtttct tatttgacgt
1201 gtactttact tgtcctaagt gtggcacaga agtacacagt ctgcagccac tgaaatcagg
1261 aatcgagatg tcagaagtaa atgctcttta gaaactaaaa ttgcttcctt tgaggaaaaa
1321 aatgcaccgt gtttaaattc ttaatattag tcatcctaag tataccattt atgtatcctt
1381 tataaggaat gtgctcctag ccactgtctt ctcttttcca ggcagtactg aaatctaata
1441 actggaaacc atgtgattct aaatatatta tgtaaagtgt aatgtattat gttttttaaa
1501 tcattgcatt caatttttaa tgtcaagaat aatggacagc ttttgtcagg tgactactaa
1561 caatgctcct tcattttact acttcttaaa acaagggttg attctaaaga taaagatttt
1621 ggagactctg
```

FIGURE 50

Figure 51

Name: Peroxisome Gen Bank Accession No. AF060502 SEQ ID No.
52
Assembly Protein 12

```
1 gtggctgctc gggaccaccc gaaccgcgg ccatggcccc ggccgcgcgc agccccccgg
61 aggtgatccg cgcggcgagc aaggacgagt actaccgcgg tgggctgcgg agcgcggcgg
121 gcggcgccct gcacagcctg gcgggtgcga ggaagtggct ggagtggagg aaggaggttg
181 agctgctctc agatgtggcc tactttggcc tcaccacact tgcaggctac cagaccctgg
241 gggaggagta cgtcagcatc atccagggtg acccatcgcg gatacatgtg cctcctcgc
301 tgcgccgtgg cgtgctggtg acgctgcatg ccgtcctgcc ctacctgctg gacaaggccc
361 tgctccccct ggagcaggag ctgcaggctg accccgacag tgggcgaccc ttgcagggga
421 gcctggggcc aggtgggcgt ggctgctcag gggcgcggcg ctggatgcgt caccacacgg
481 ccaccctgac tgagcagcag aggaggggcg tgetgcgggc ggtcttcgtc ctcagacagg
541 gcctcgctg cctccagcgg ctacatgttg cctggtttta catccacggt gtcttctacc
601 acctggccaa gaggtcacg gggatcacgt acctccgtgt ccgcagcctg cccggagagg
661 acctgagggc ccgtgttagc tacaggctgc tgggggtcat ctactgctg cacctggtgc
721 tgtccatggg gctgcagctg tacggtttca ggcagcggca gcgagccagg aaggagtgga
781 ggctgcaccg cggcctgtct caccgcaggg cctccttggg ggagagagcc gtttccagaa
841 accccctgtg caccctgtgc ctggaggagc gcaggcacc aacagccacg cctgcgggc
901 acctgttctg ctgggagtgc atcaccgctt ggtgcagcag caaggcggag tgtcccctct
961 gccgggagaa gttcccctcc cagaagctca tctaccttcg gcactaccgc tgagccggcg
1021 cccgggtggg cctggacaca gatgacctct acgggagtct aaacgccaa atttagtctc
1081 aggattaacc ttgcttgac agaagttaga acactctcag tttttgtca tgtaagatac
1141 taacctagcc accctgggag agaacagaaa gctgtccctg gctgcgcttt ctacgcctg
1201 ggagggggcg ctgaaccag aacatttccc taaccccaac ctggtaggac tcagccactt
1261 cttcagg
```

FIGURE 51

Figure 52

Name: Peroxisome

Gen Bank Accession No. U91521

SEQ ID No. 53

Assembly Protein 12

```

1  cgccccgggca ggtctccctt caaacgtggc gtcgtgggtt gtttgcgcct cgccctggggg
61  cagcgagcaa ggacggggcg gggcggggat actcaaagcc aacagctgga gtcagccctc
121  gtgtcccggg ctcacagtgg cagcactgaa tcctcagagt cggctggctt ttgagctctc
181  acgattgggg aggagggggc gttcttggtt cgcagctcca gaggattgag ttccctcccc
241  catacctgtc cccacagtc acgctctgcc ctgacgtgca gcatttgaca agttaccccc
301  tcgccacata ctacttcac ccacgtccga gttaactttg ttcttaacct tcttgagact
361  accctcggcc tccaggtctt tttttccag ttcatttttg ccataagat tgagtttcga
421  gtttcagata tcatgcagaa agtttacctt taagactgag caccatctg atactcttcc
481  tccccaaaaa gttcatgtc acgagagagt ttgtgggaaa agtgaaagcc agtacacgca
541  ggaaactatg gctgagcacg gggctcactt cacagctgct tctgtggccg atgaccagcc
601  atccatcttt gagtggttag cacaggacag tttaatgaca gcagtggagc ccgctcttca
661  gcatgtggtc aaggttcttg cagaatcaaa tcccaccac tatggcttct tgtggaggtg
721  gtttgatgaa atctttactc tgctagatct tctgtccag caacattatc tgtctagaac
781  cagtgcctca tttctgaaa acttttacgg cttaaagaga attgtaatgg gggacactca
841  caagtctcag agattggcta gtgctggtct cccaaagcag cagctttgga aatctattat
901  gttcctggtt cttctccct atctgaaagt gaagctggag aagctggttt ctagcctgag
961  agaagaggat gaatattcta ttcaccccc ttctcccg cggaaacgat ttacagagc
1021  tttcctggca gcctacccat ttgtgaacat ggcctgggaa ggatggtttc ttgtacaaca
1081  acttcgatac atcctaggaa aagctcagca tcactacca ctgctgaggc tggctggagt
1141  tcagctaggt cgactgacag ttcaggatat acaagctctg gagcacaac cagctaaggc
1201  cagcatgatg cagcaaccag ccaggagtgt tagtgagaag ataaactcag ctctgaagaa
1261  agctgttggg ggtgttgctt tatccctgtc tactggcctt tctgtgggtg tattcttctt
1321  gcagttcctt gactggtggt actcatctga aaatcaagaa accatcaagt cattgactgc
1381  cctgcctact ccaccaccac ctgtacacct agactataac tctgattctc cctcttacc
1441  caaaatgaag actgtgtgcc cactgtgtcg taaaaccgg gtgaatgata ctgttcttgc
1501  cacctctggc tatgtgtttt gttaccgctg tgtgtttcat tatgtgagga gtcaccaagc
1561  ttgtcccatc acaggttatc caacagaagt acaacatctg attaaactct actccctga
1621  gaactgaaag ggaatcatgt cttatcctca caacaaaagt attgcgctgt aatgtcacag
1681  agctggtttc agcatggtat tgagttgaca cttctcagcc tcaattcata actatatgaa
1741  ctttaaataa ttatatggat ttctttaaaa ggatatgcc atttgatcct aaccttttag
1801  cctgctctct agaaccctac ttacagttga ccaagctgct taaagtagaa ggatcaggac
1861  tggcagggga tatgagggcc agtgtagaga ttaaacagaa taatgagtac acaagttttt
1921  gcctatccct ttctttaaaa tataagaaat aaaggagctc ttacgaaaaa cccaggttag
1981  accgcataaa aaataaagtg aacagtgagg tggtagcaag acttcttttt agaaaagaaa
2041  gcatttacct gcctgtctgt aaggtggaaa tttcatcagt ttgcaaacga taagaaatgc
2101  agacttgctc ttgatagaaa tgcttagaaa cactctgggg gggaaaaaag ctcttccat
2161  atactgtgag acatttggtt agtgacatct attgtttatc agcttttaag gataaaaaag
2221  gtatttttaa agttggatat ttaggatatt tgaggatatt ctttatgag ctctccatat
2281  cttcttgag aaactgggtt aaaaaggaat aggggtagag tgttacagag agtagtctga
2341  agattcctgt gtaaaagcaa agctaacaag caatgaagac atgaagcaaa atactaatct
2401  aattgtgtaa agaaaggata ttttaataag ttctttctgc ttgctgctaa gagtttgcta
2461  aagtgtcatg aattattctg gttattacta aagtttctat gaaacaagta gattttaaga
2521  ataaatgttt ctggaaaagg actatgttat gattttgtag aaatgtaaag attacttgag
2581  gtgtttaaaa taaatttttc attcagact

```

FIGURE 52

Figure 53

Name: AcetylcholineGen Bank Accession No. Z33905SEQID No. 54

Receptor-Associated

43 KDA Protein

```
1  cccaactggc agcgacacgt agggacgggc tgaaccagct ttcttcccag ggtggcgccct
61  gctctccatc caggcccat tccggctccc acccgactgc tgcttttgtt cccacgtttc
121  ggggggcagc tggcactgtg attcctgccc catgagtgcc tagtggccag gagccaccag
181  ggatcacccc acgtgggctt ggggaggatg gggcaggacc agaccaagca gcagatcgag
241  aaggggctcc agctgtacca gtccaaccag acagagaagg cattgcaggt gtggacaaag
301  gtgctggaga agagctcgga cctcatgggg cgcttccgcg tgetgggctg cctggtcaca
361  gccactcgg agatgggccc ctacaaggag atgctgaagt tcgccgtggt ccagatcgac
421  acggcccggg agctggagga tgccgacttc ctctggaga gctacctgaa cctggcacgc
481  agcaacgaga agctgtgcga gtttcacaag accatctcct actgcaagac ctgccttggc
541  gtgccaggta ccagggcagg tgcccagctc ggaggccagg tcagcctgag catgggcaat
601  gccttcttgg gcctcagcgt cttccagaag gccctggaga gcttcgagaa ggccctgcgc
661  tacgcgcaca acaatgatga caccatgctc gagtgccgcg tgtgctgcag cctgggcagc
721  ttctatgccc aggtcaagga ctacgagaaa gccctgttct tcccctgcaa ggcggcagag
781  cttgtcaaca actatggcaa aggtctggagc ctgaagtacc aggccatgag ccagtaccac
841  atggccgtgg cctatgcctt gctgggcccgc ctgggcagtg ccatggagtg ttgtgaggag
901  tctatgaaga tcgccctgca gcacggcgac cggccactgc aggcgctctg cctgctctgc
961  ttcgctgaca tccaccggag ccgtggggac ctggagacag ccttccccag gtacgactcc
1021 gccatgagca tcatgaccga gatcggaaac cgctggggc aggtgcagge gctgctgggt
1081 gtggccaagt gctgggtggc caggaaggcg ctggacaagg ctctggatgc catcgagaga
1141 gcccaggatc tggccgagga ggtggggaac aagctgagcc agctcaagct gcaactgtctg
1201 agcgagagca ttaccgcag caaagggtcg cagcgggaac tgcgcgcgca cgttgtgagg
1261 ttccacgagt gcgtggagga gacggagctc tactgcgcgc tgtgcggcga gtccataggc
1321 gagaagaaca gccggctgca ggccctaccc tgctcccaca tcttccacct caggtgcctg
1381 cagaacaacg ggaccgggag ctgtcccaac tgccgccgct cctccatgaa gcctggcttt
1441 gtatgactcc tggcaggagg gcgtgggctt cctcctgggc gactcctgct ctttctccac
1501 tgcaccgagt ggccattact cctgggcagc tgccaggtct gcctcaccat acgcaacgct
1561 tggggccccg agggcgtgct cccctggccc agctcccctg ccctgcctct ttgtactttg
1621 ctctttatag aaaaataaac tgtttgtacc tgg
```

FIGURE 53

Figure 54

Name: Peroxisomal Gen Bank Accession No. BC000661 SEQ ID No.
55
Membrane Protein 3

```
1  ggcacgaggg aaaagacaag ctttccagag aatgtttcag agaaagttac gtggagcgtg
61  ggcgttttcgc agactcctaa gtggtctgga acctaacctg cagtgtctct gagcttctgg
121 caagtatgag gcaaggattt cacagaagaa cttggagcaa aaattcccac tgtcatagct
181 ataggctgtc cagtcttgag ggagctgcag gaacaggaaa aagagacctt cagagaagac
241 atggccttcca gaaaagagaa tgcgaagagt gcaaacagag tgctaagaat aagccagttg
301 gatgcacttg aactaaacaa ggccctggag cagctagttt ggtcccagtt tactcagtg
361 tttcatggat ttaaacctgg gctgttagct cgctttgagc cagaggtgaa agcgtgctta
421 tgggttttct tgtggagatt caccatctac tccaaaaatg ccacagtggg acagtcagtt
481 ttgaatatta agtacaaaaa tgatttttcc cctaacctga gatatcagcc acccagtaaa
541 aatcaaaaaa tctggtatgc tgtttgtaca attggtggca ggtgggttaga agaacgatgc
601 tatgatttgt ttcgaaacca tcatttagca tcatttggga aagtcaagca gtgtgtgaat
661 tttgtgattg gacttttgaa attaggtggg ctgattaatt ttttgatttt ccttcagagg
721 ggaaagtttg caactttgac agaacgtctc ctaggtattc attctgtatt ttgcaagcct
781 caaaacatac gtgaagttgg ctttgaatac atgaataggg aacttctctg gcatggtttt
841 gctgaatttc tgatttttct cttaccactt atcaatgtcc agaagttgaa agccaagctg
901 tcttcattgg gtattcctct tactggtgca cctaatagtg acaatacatt agccaccagt
961 ggcaaagaat gcgctctatg tggagagcgg cccaccatgc ctcacaccat aggatgtgag
1021 catattttct gttatttctg tgctaagagt agtttcttat ttgacgtgta ctttacttgt
1081 cctaagtgtg gcacagaagt acacagtctg cagccactga aatcaggaat cgagatgtca
1141 gaagtaaata ctcttttaga actaaaattg cttcctttga ggaaaaaaat gcaccgtgtt
1201 taaattctta atattagtca tcctaagtat accatttatg tatcctttat aaggaatgtg
1261 ctcttagcca ctgtcttctc ctttccaggc atgactgaaa tctaatactt ggaaaccatg
1321 tgattctaaa tatattatgt aaatgttaat gtattatgtt ttttaaatca ttgcattcaa
1381 tttttaatgt caagaataat ggacagcttt tgtcaggtga ctactaacia tgctccttca
1441 ttttactact tcttaaaaca aggttggatt ctaaagataa agattttgga gactctgaaa
1501 aaaaaaaaaa aaaaaaaaaa aaaaaaa
```

FIGURE 54

Figure 55

Name: HsRmal
56

Gen Bank Accession No. AB056869

SEQ ID No.

```
1 atggcagcag cggaggagga ggacgggggc cccgaagggc caaatcgca gcggggcggg
61 gcgggcgca ccttcgaatg taatatatgt ttggagactg ctcgggaagc tgtggtcagt
121 gtgtgtggcc acctgtactg ttggccatgt cttcatcagt ggctggagac acggccagaa
181 cggcaagagt gtccagtatg taaagctggg atcagcagag agaaggttgt cccgctttat
241 gggcgagggg gccagaagcc ccaggatccc agattaaaaa ctccaccccg ccccagggc
301 cagagaccag ctccggagag cagaggggga ttccagccat ttggtgatac cgggggcttc
361 cacttctcat ttggtgttgg tgcttttccc tttggctttt tcaccaccgt cttcaatgcc
421 catgagcctt tccgccgggg tacagggtgtg gatctgggac agggtcaccc agcctccagc
481 tggcaggatt ccctcttctt gtttctcgcc atcttcttct ttttttggct gtcagtatt
541 tga
```

FIGURE 55

Figure 56

Name: HTRIP
57

Gen Bank Accession No. U77845

SEQ ID No.

```
1 gtgcggtgga gcgaaatttg aagcaagcgg aggcggggcg ctctacgaag ccggacctgt
61 agcagtttct ttggctgcct gggccccttg agtccagcca tcatgcctat cgtgctctg
121 tgcactatct gctccgactt cttcgatcac tcccgcgacg tggccgccat ccactgcggc
181 cacaccttcc acttgacgtg cctaattcag tcctttgaga cagaccaag tcggacctgc
241 ccacagtgcc gaatccaggt tggcaaaaga accattatca ataagctctt ctttgatctt
301 gccagggagg aggagaatgt cttggatcga gaattcttaa agaatgaact ggacaatgtc
361 agagcccagc tttcccagaa agacaaggag aaacgagaca gccaggtcat catcgacact
421 ctgcgggata cgctggaaga acgcaatgct actgtggtat ctctgcagca ggccctgggc
481 aaggccgaga tgctgtgctc cactctgaaa aagcagatga agtacttaga gcagcagcag
541 gatgagacca aacaagcaca agaggaggcg ggccggctca ggagcaagat gaagaccatg
601 gagcagattg agcttctact ccagagccag ctccctgagg tggaggagat gatccgagac
661 atgggtgtgg gacagtcagc ggtggaacag ctggctgtgt actgtgtgtc tctcaagaaa
721 gagtacgaga atctaaaaga ggcacggaag gcctcagggg aggtggctga caagctgagg
781 aaggatttgt tttcctccag aagcaagttg cagacagtct actctgaatt ggatcaggcc
841 aagttagaac tgaagtcagc ccagaaggac ttacagagtg ctgacaagga aatcatgagc
901 ctgaaaaaga agctaacgat gctgcaggaa accttgaacc tgccaccagt ggccagttag
961 actgtcgacc gcctggtttt agagagccca gccctgtgtg aggtgaatct gaagctccgc
1021 cggccatcct tccgtgatga tattgatctc aatgctacct ttgatgtgga tactcccca
1081 gcccgccct ccagctcca gcatggttac tacgaaaaac tttgcctaga gaagtcacac
1141 tccccaatc aggatgtccc caagaagata tgcaaaggcc ccaggaagga gtcccagctc
1201 tcaactgggtg gccagagctg tgcaggagag ccagatgagg aactggttg tgccctccct
1261 atttttgtcc ggaatgccat cctaggccag aaacagccca aaaggcccag gtcagagtcc
1321 tcttgacgca aagatgtggt aaggacaggc ttcgatgggc tcggtggccg gacaaaattc
1381 atccagccta ctgacacagt catgatccgc ccattgcctg ttaagcccaa gaccaaggtt
1441 aagcagaggg tgaggtgaa gaccgtgcct tctctcttcc aggccaagct ggacaccttc
1501 ctgtggtcgt gagaacagtg agtctgacca atggccagac acatgcctgc aacttgtagg
1561 tcaaggactg tccaggcagg gtttgtggac agagccctac tttcgggacc agcctgaggt
1621 gtaagggcag acaaacaggt gaggggtgagt gtgacacca gagactgctc ttctgcctt
1681 caccctgccc cactctacg actgggagct gacatgacca gccactgat cctgtcagca
1741 ggtcctgctc tgttgccagg ctcttgttta tagccatgat cagatgtggt cagactcttt
1801 ctgggccttg agaccacggt cacttgttga ctgtctctgt ggaccagagt gcttgaggca
1861 tctcaggcag cctcagcca agcttctacc tgcccttgac ttgcttctag catagcctgg
1921 gccaaagcagg gtggggaatg gaggatagac atgggatgta tggagaggat ggaagatttt
1981 cccgaaaaaa aaaaaaaaa aaaaaaa
```

FIGURE 56

Figure 57

Name: Cystein Rich Domain GenBank Acc.# X80200 SEQ ID No. 58
 Associated to RING
 and TRAF Protein

```

1 gccgggagcg ccgctccagc gaggcgcggg ctgtggggcc gccgcgtgcc tggccccgct
61 cgcccgtgcc ggccgctcgc ccgccatgcc tggttcgac tacaagttcc tggagaagcc
121 caagcgacgg ctgctgtgcc cactgtgcgg gaagcccatg cgcgagcctg tgcaggtttc
181 cacctgcggc caccgtttct gcgatacctg cctgcaggag ttccctcagt aaggagtctt
241 caagtgcctt gaggaccagc ttccctctga ctatgccaa atctaccag acccgagct
301 ggaagtacaa gtattggggc tgcctatccg ctgcatccac agtgaggagg gctgccgctg
361 gagtggggcca ctacgtcatc tacaggggcca cctgaatacc tgcagcttca atgtcattcc
421 ctgccctaata cgtgccccca tgaagctgag ccgccgtgat ctacctgcac acttgacgca
481 tgactgcccc aagcggcgcc tcaagtgcga gttttgtggc tgtgacttca gtggggaggc
541 ctatgagagc catgagggtg tgtgccccca ggagagtgtc tactgtgaga ataagtgtgg
601 tgcccgcagt atgcgggggc tgctggccca gcatgccacc totgagtgcc ccaagcgcac
661 tcagccctgc acctactgca ctaaggagtt cgtctttgac accatccaga gccaccagta
721 ccagtgcccc aggtgcctg ttgcctgccc caaccaatgt ggtgtgggca ctgtggctcg
781 ggaggacctg ccaggccatc tgaaggacag ctgtaacacc gccctggtgc tctgcccatt
841 caaagactcc ggctgcaagc acagggtgcc taagctggca atggcacggc atgtggagga
901 gagtgtgaag ccacatctgg ccatgatgtg tgccctggtg agccggcaac ggcaggagct
961 gcaggagctt cggcgagagc tggaggagct atcagtgggc agtgatggcg tgctcatctg
1021 gaagattggc agctatggac ggcggctaca ggaggccaag gccaagcca acctgagtg
1081 cttcagccca gccttctaca cacataagta tggttacaag ctgcagggtg ctgcattcct
1141 caatggcaat ggcagtgggt agggcacaca cctctcactg tacattcgtg tgctgcctgg
1201 tgcccttgac aatctccttg agtggccctt tgcccgcctg gtcaccttct cctgctgga
1261 tcagagcgac cctgggctgg ctaaaccaca gcacgtcact gagaccttcc accccgacce
1321 aaactggaag aatttccaga agccaggcac gtggcggggc tccttgatg agagtctct
1381 gggctttggt tatcccaagt tcctctcca ccaggacatt cgaaagcgaa actatgtgcg
1441 ggatgatgca gtcttcatcc gtgctgctgt tgaactgcc cggaagatcc tcagctgagt
1501 gcagggtggg ttcgagggga aaggacgatg gggcatgacc tcagtcaggc actggctgaa
1561 cttggagagg gggccggacc ccctcagct gcttctgctg cctaggttct gttaccccat
1621 cctccctccc ccagccacca ccctcaggtg cctccaattg gtgcttcage cctggccctt
1681 gtggggaaca ggtcttgggg tcatgaaggg ctggaaacaa gtgacccag ggctgtctc
1741 ctttcttggg tagggcagac atgccttggg gccggtcaca ctctacacgg actgaggtgc
1801 ctgctcaggt gctatgtccc aagagccata aggggtggg aattggggag ggagaaaggg
1861 tagttcaaag agtctgtctt gagatctgat ttttcccc tttacctagc tgtgcccct
1921 ctgggttatt attccttag tgccaggagg gcacagcagg ggagccctga tttttaataa
1981 atccggaatt gtatttatt

```

FIGURE 57

Figure 58

Name: PML-RARGen Bank Accession No. M73779SEQ ID No.59

1 ctcccccttca gcttctcttc acgcactcca agatctaaac cgagaatcga aactaagctg
 61 ggggtccatgg agcctgcacc cgcccgatct ccgaggcccc agcaggaccc cgcccgcccc
 121 caggagccca ccatgcctcc ccccgagacc cctctgaag gccgccagcc cagccccagc
 181 cccagcccta cagagcgagc ccccgcttcg gaggaggagt tccagtttct gcgctgccag
 241 caatgccagg cggaagccaa gtgccgaag ctgctgcctt gtctgcacac gctgtgctca
 301 ggatgcctgg aggcgtcggg catgcagtgc cccatctgcc aggcgccctg gcccctaggt
 361 gcagacacac ccgccctgga taacgtcttt ttcgagagtc tgcagcggcg cctgtcgggtg
 421 taccggcaga ttgtggatgc gcaggctgtg tgcacccgct gcaaagagtc ggccgacttc
 481 tgggtgctttg agtgcgagca gctcctctgc gccaaagtgt tgcaggcaca ccagtgggtc
 541 ctcaagcacg agggccggcc cctagcagag ctgcgcaacc agtcgggtgc tgagttcctg
 601 gacggcaccg gcaagaccaa caacatcttc tgtcccaacc ccaaccaccg caccctacg
 661 ctgaccagca tctactgccg aggatgttcc aagccgctgt gctgctcgtg cgcgtcctt
 721 gacagcagcc acagtgcgt caagtgcgac atcagcgcag agatccagca gcgacaggag
 781 gagctggacg ccatgacgca ggcgtgcag gagcaggata gtgcctttgg cgcggttcac
 841 gcgcagatgc acgcggccgt cggccagctg ggccgcgcgc gtgccgagac cgaggagctg
 901 atccgcgagc gcgtgcgcca ggtggtagct cacgtgcggg ctgaggagcg cgagctgctg
 961 gaggctgtgg acgcgcggta ccagcgcgac tacgaggaga tggccagtcg gctgggcccg
 1021 ctggatgctg tgctgcagcg catccgcacg ggcagcgcgc tgggtgcagag gatgaagtgc
 1081 tacgcctcgg accaggaggt gctggacatg cacggtttcc tgcgccaggc gctctgccgc
 1141 ctgcgccagg aggagcccca gagcctgcaa gctgccgtgc gcaccgatgg ctctgcagag
 1201 ttcaaggtgc gcctgcagga cctcagctct tgcacaccc aggggaaagc cattgagacc
 1261 cagagcagca gttctgaaga gatagtccc agccctccct cgccacccc tctaccccgc
 1321 atctacaagc cttgctttgt ctgtcaggag aagtcctcag gctaccacta tggggtcagc
 1381 gctgtgagg gctgcaagg cttcttcgc cgcagcatcc agaagaacat ggtgtacacg
 1441 tgtcacccgg acaagaactg catcatcaac aaggtgacct ggaaccgctg ccagtactgc
 1501 cgaactgcaga agtgctttga agtgggcagt tccaaggagt ctgtgagaaa cgaccgaaac
 1561 aagaagaaga aggaggtgcc caagcccag tgccttgaga gctacacgct gacgccggag
 1621 gtgggggagc tcattgagaa ggtgcgcaaa gcgcaccagg aaaccttccc tgcctctgc
 1681 cagctgggca aatacactac gaacaacagc tcagaacaac gtgtctctct ggacattgac
 1741 ctctgggaca agttcagtga actctccacc aagtgcacat ttaagactgt ggagttcgcc
 1801 aagcagctgc ccggcttcac caccctcacc atcgccgacc agatcaccct cctcaaggct
 1861 gcctgcctgg acatcctgat cctgcggatc tgcacgcggt acacgcccga gcaggacacc
 1921 atgaccttct cggacgggct gacctgaac cggaccaga tgcacaacgc tgggttcggc
 1981 cccctcaccg acctggtctt tgcttcgcc aaccagctgc tgccttgga gatggatgat
 2041 gcggagacgg ggtgctcag cgcctctgc ctcatctgc gagaccgcca ggacctggag
 2101 cagccggacc ggggtgacat gctgcaggag ccgctgctgg aggcgctaaa ggtctacgtg
 2161 cggaagcgga ggcccagccg cccccacatg ttccccaaaga tgctaataa gattactgac
 2221 ctgcgaagca tcagcgccaa gggggtgag cgggtgatca cgctgaagat ggagatcccg
 2281 ggctccatgc cgctctcat ccaggaaatg ttggagaact cagagggcct ggacactctg
 2341 agcggacagc cggggggtgg ggggcgggac gggggtggcc tggcccccgc gccaggcagc
 2401 tgtagcccca gcctcagccc cagctccaac agaagcagcc cggccaccca ctccccgtga
 2461 ccgcccacgc cacatggaca cagccctcgc cctccgcccc ggcttttctc tgcttttcta
 2521 ccgaccatgt gacccgcac cagccctgcc cccacctgcc ctcccgggca gtactgggga
 2581 ccttccctgg gggacgggga gggaggaggc agcgactcct tggacagagg cctgggccct
 2641 cagtggactg cctgctccca cagcctgggc tgacgtcaga ggccgaggcc aggaactgag
 2701 tgaggccctt ggtcctgggt ctgaggatgg gtccctgggg cctcgtgttc atcaagacac
 2761 cctctgccc agctcaccac atcttcatca ccagcaaac ccaggacttg gctcccccat
 2821 cctcagaact cacaagccat tgcctcccag ctgggggaacc tcaacctccc cctgcctcg
 2881 gttggtgaca gggggggtgg gacaggggag gggggttccc cctgtacata cctgccata
 2941 ccaacccag gtattaatc tcgtggttt tgtttttatt ttaatttttt tgttttgatt
 3001 tttttaataa gaattttcat ttttaagcaaa aaaaaa

FIGURE 58

Figure 59

Name: TRAF5Gen Bank Accession No. AB000509SEQID No. 60

```
1 gcagcagccg cgccctgcaga ccggcctcgc ggagcccgcg cgccgagccc cacaatggct
61 tattcagaag agcataaagg tatgcctgtt ggtttcatcc gccagaattc cggcaactcc
121 atttccttgg actttgagcc cagtatagag taccagtttg tggagcgggt ggaagagcgc
181 tacaaatgtg cctttctgcca ctcggtgctt cacaaccccc accagacagg atgtgggcac
241 cgcttctgcc agcactgcat cctgtccctg agagaattaa acacagtgcc aatctgccct
301 gtagataaag aggtcatcaa atctcaggag gtttttaaag acaattgttg caaaagagaa
361 gtcctcaact tatatgtata ttgcagcaat gctcctggat gtaatgccaa ggttattctg
421 ggccggtacc aggtacacct tcagcagtgc ttatttcaac ctgtgcagtg ttctaagag
481 aagtgccggg agccagtcct acggaaagac ctgaaagagc atttgagtgc atcctgtcag
541 tttcgaaagg aaaaatgcct ttattgcaaa aggatgtgg tagtcatcaa tctacagaat
601 catgaggaaa acttgtgtcc tgaatacca gtattttgtc ccaacaattg tgcgaagatt
661 attctaaaaa ctgaggtaga tgaacacctg gctgtatgtc ctgaagctga gcaagactgt
721 ccttttaagc actatggctg tgcgtgaacg gataaacgga ggaacctgca gcaacatgag
781 cattcagcct tacgggagca catgcgtttg gttttagaaa agaattgtcc attagaagaa
841 cagatttctg acttacacaa gagcctagaa cagaaagaaa gtaaaatcca gcagctagca
901 gaaactataa agaaacttga aaaggagtgc aagcagtttg cacagttgtt tggcaaaaat
961 ggaagcttcc tcccaaacat ccagggtttt gccagtcaca ttgacaagtc agcttggcta
1021 gaagctcaag tgcattcaatt attacaaatg gtttaaccagc aacaaaataa atttgacctg
1081 agacctttga tggaagcagt tgatacagtg aaacagaaaa ttacctgtct agaaaacaat
1141 gatcaaagat tagccgtttt agaagaggaa actaacaaac atgataccca cattaatatt
1201 cataaagcac agctgagtaa aaatgaagag cgattttaa acgtgtggagg tacttgctat
1261 aatggaaagc tcatttggaa ggtgacagat tacaagatga agaagagaga ggcggtggat
1321 gggcacacag tgtccatctt cagccagtc ttctacacca gccgctgttg ctaccggctc
1381 tgtgctagag catacctgaa tggggatggg tcagggaggg ggtcacacct gtccctatac
1441 tttgtggtca tgcgaggaga gtttgactca ctgttgagc ggccattcag gcagagggtg
1501 accctgatgc ttctggacca gagtggcaaa aagaacatta tggagacctt caaacctgac
1561 cccaatagca gcagctttaa aagacctgat ggggagatga acattgcatc tggctgtccc
1621 cgctttgttg ctcattctgt tttggagaat gccaaagac cctacattaa agatgacact
1681 ctgttcttga aagtggccgt ggacttaact gacctggagg atctctagtc actgttatgg
1741 ggtgataaga ggacttcttg gggccagaac tgtggaggag agcacatttg attatcatat
1801 tgacctggat ttagactcaa agcacatttg tatttgcctt tttccttaac gtttgaagtc
1861 agtttaaaac ttctgaagtg ctgtcttttt acattttact ctgtcccagt ttgaaactta
1921 aaactcttag aatattctct tattatttat atttttatat ttcttgaaag atggtaagtt
1981 tcttgaagtt tttggggcgt ttctctttta ctgggtgctta gcgcagtgtc tcgggcactc
2041 taaatattga gtgttatgga ggacacagag gtagcagaat ccagttgaa aatgttttga
2101 tattttattg tttggcctat tgattctaga cctggcctta agtctgcaaa agccatcttt
2161 ataaggtagg ctgttccagt taagaagtgg gtgatgtagt taaaaagata atatgctcag
2221 tttggacctt tttttcagtt aaatgctaaa tatatgaaaa ttactatacc tctaagtatt
2281 ttcatgaaat tcaccagcag tttgcaagca cagttttgca aggctgcata agaactggtg
2341 aatggggtaa gcattttcat tcttctgtct gaagtaaagc agaaagtact gcatagtata
2401 tgagatatag ccagctagct aaagttcaga ttttgttagg ttcaacccta tgaaaaaac
2461 tattttcata ggtcaaaaat ggtaaaaaat tagcagtttc ataagattca accaaaataa
2521 tatatatata cacacacaca tacatatata cctatatatg tgtgtatata aacagttcga
2581 atgtattttg gtgacagtaa taaatcaatg tgaggatgga tagaatttag tatatgatag
2641 agaaaatgtc ataaatggat aaaaggaatt tacaacttga ggagaaaacc tttacaattt
2701 cctatgggtg tcagaagtac tctcagcgaa aactgatggc taaaacagta tctactattc
2761 tctgataact ttttttttga gacagagttt cattgtcacc caggctggag tacagtggca
2821 tgatctcagc tctactgcaa ctctgcctcc cgaattcaag tgattctcct gcctcagcct
2881 cctgagtagc tgggattaca ggcccccgtc accacaccca ggtaattttt gtatttttag
2941 tagagacgga gttttgccat gttggccaag ctgatctcaa actcctgacc tcaagtgatc
3001 tgcccgcctc ggcctcccaa agtgctgaga ttacaggcat gaccaccgc gtcaagcctc
3061 tgacaactat tgaatttgta agctgctatg caaatgggca tttatataaa cttgtgatgt
3121 ttcttgtcag aattctgagt actctgtgaa gaacagaaat gatcatattc ttatgcatct
```

```
3181 atctgtatgg gtctgaaggt gtatatatacaa actgagatga gtccttatga ctcttgataa
3241 gcctgagttt aacaacaaca aaaatgccaa gttgtcctga gcccttctgc gttgttatgc
3301 cacttcccta ctgctcataat gcacgctggc tcccctgggc acgcaaggat gagtatgggc
3361 catgggcccc tgtagagctg cttacctggg gatgaccatg caccttaca tttctgaaca
3421 gtttaacccta tagaagcatg ctttatatga gtgtcttctg ggaagaggaa ccttcttaat
3481 ctcttctgtg ggattttcaa aatgctaaag actcacactg cagcaatcat ccagatgat
3541 taaattcaaa gaaatagggt cacaacagga atatactgaa gaactagagt gtcactgctg
3601 gtgaactgtg gcacggttgc tcaacacatc acctcggaca aattcaggaa gcatttcttt
3661 agcccacaag tccagacca ggtgctctgt atgtttgttt ttaatatcca tcatatccaa
3721 gttcactctg tcttctctgag cagtggaga tcatattgct gtaacttctt ttaagtagtt
3781 gatgtggaaa acattttaaa gtgaatttgt caaatgctg gttttgtgtt ttatccaact
3841 tttgtgcata tatataaagt atgtcatggc atggtttgct taggagttca gagttccttc
3901 atcatcgaaa tagtgattaa gtgatcccag aacaaggaat actagagtaa aaagcacctc
3961 tttttcagaa aaaaaaaaaa aaaaaaaaaa aaa
```

FIGURE 59

Figure 60

Name: Tumor Necrosis Gen Bank Accession No. AF082185SEQ ID No.61

Factor Receptor-Associated

Factor 4A

```
1 atgcctggct tegactacaa gttcctggag aagcccaagc gacggctgct gtgcccactg
61 tgcgggaagc ccatgcgcga gcctgtgcag gtttccacct gcggccaccg tttctgcgat
121 acctgcctgc aggagttcct cagtgaagga gtcttcaagt gccctgagga ccagcttcct
181 ctggactatg ccaagatcta cccagacccg gagctggaag tacaagtatt gggcctgcct
241 atccgctgca tccacagtga ggagggctgc cgctggagtg ggccactacg tcatctacag
301 ggccacctga atacctgcag cttcaatgtc attccctgcc ctaatcgctg ccccatgaag
361 ctgagccgcc gtgatctacc tgcacacttg cagcatgact gcccgaagcg gcgcctcaag
421 tgcgagtttt gtggctgtga cttcagtggtg gaggcctatg agagccatga gagttctctg
481 ggctttggtt atcccaagtt catctccac caggacattc gaaagcgaaa ctatgtgcgg
541 gatgatgcag tcttcatccg tgctgctggt gaactgcccc ggaagatcct cagctga
```

FIGURE 60

Figure 61

Name: HFB30Gen Bank Accession No. AB022663SEQID No. 62

```

1 gaagaggcgg cgcgccggtt gaggaggcg tttctaggcc tggtcggctg gggcgatgg
61 caggattttc ataatatatg tagtatgagt tccacatctt ggctcttac ccagcttcag
121 cagtctcagc tccaccagtt agagaataaa tgggatttgc atgaactcca ctctgagctg
181 aaatacagac tgcggtgtta atatcctttt cttctgattg ttattaacag gtccctatgt
241 cgtcagaaga tcgagaagct caggaggatg aattgctggc cctggcaagt atttacgatg
301 gagatgaatt tagaaaagca gagtctgtcc aagggtggaga aaccaggatc tatttggatt
361 tgccacagaa tttcaagata tttgtgagcg gcaattcaaa tgagtgtctc cagaatagtg
421 gctttgaata caccatttgc tttctgcctc cacttgtgct gaactttgaa ctgccaccag
481 attatccatc ctcttcccca ccttcattca cacttagtgg caaatggctg tcaccaactc
541 agctatctgc tctatgcaag cacttagaca acctatggga agaaccctg ggcagcgtgg
601 tcctgtttgc ctggatgcaa tttcttaagg aagagaccct agcatacttg aatattgtct
661 ctccctttga gctcaagatt ggttctcaga aaaaagtgcg gagaaggaca gctcaagctt
721 ctccaacac agagctagat tttggaggag ctgctggatc tgatgtagac caagaggaaa
781 ttgtggatga gagagcagtg caggatgtgg aatcactgtc aaatctgatc caggaaatct
841 tggactttga tcaagctcag cagataaaat gctttaatag taaattgttc ctgtgcagta
901 tctgtttctg tgagaagctg ggtagtgaat gcatgtactt cttggagtgc aggcattgtg
961 actgcaaagc ctgtctgaag gactactttg aaatccagat cagagatggc caggttcaat
1021 gcctcaactg cccagaacca aagtgcctt cggtggccac tcctggtcag gtcaaagagt
1081 tagtgggaagc agagttatatt gcccgttatg accgccttct cctccagtc tccctggacc
1141 tgatggcaga tgtggtgtac tgcccccggc cgtgctgccg gctgcctgtg atgcaggaac
1201 ctggctgcac catgggtatc tgctccagct gcaattttgc cttctgtact ttgtgcaggt
1261 tgacctacca tggggtctcc ccagtgaagg tgactgcaga gaaattaatg gacttacgaa
1321 atgaatacct gcaagcggat gaggctaata aaagactttt ggatcaaagg tatggttaaga
1381 gagtgattca gaaggcactg gaagagatgg aaagtaagga gtggctagag aagaactcaa
1441 agagctgccc atgttgtgga actcccatag agaaattaga cggatgtaac aagatgacat
1501 gtactggctg tatgcaatat ttctgttggg tttgcatggg ttctctctct agagcaaac
1561 cttacaaaca tttcaatgac cctggttcac catgttttaa ccggctgttt tatgctgtgg
1621 atgttgacga cgatatattg gaagatgagg tagaagacta gttaactact gctcaagata
1681 tggaagtgga ttgtttttcc ctaatcttcc gtcaagtaca caaagtaact ttgcccagata
1741 tttagggtac tattcattca ctcttctgct gtagaagata tggagaagc aggtttatat
1801 tttcatgtgg tactactgaa gaaggtgcat tgatacattt ttaaattgaa gttgagaaaa
1861 atttataagc caaaggttca gaaaattaaa ctacagaata ttaaattatta taatgtgccc
1921 aaagctctga atagttaaaa attaaatatt tattttcttc cccaagcttt aggttaaggag
1981 aagaggggtc aagagttaaa cttagagacc ctttgtctct gagaagcatc cttotaagac
2041 attctgttgg agttccctca gtactattcc ttacaactgg agtgggtaga agccttatga
2101 aaattatact gacaacctga tctcgtttac tccatgttaa tcacattcct accaacctaa
2161 atttctgcca agtcagttct ctttggagga agcccattac tccataactg actggatggg
2221 ccagtgtcat tttgatctgc ttttcagaat ggaaatttat aatataaata tatgttttaa
2281 taccacaaac actgtggggc acttgatatt tattagtggg gtatatggtc cactggtttc
2341 aggccaaatc tagaatttag tgatactggc tcagaagaat ttaagttcta ttcagctttc
2401 ctgggtgtta gaacctagat tcaaaatggc ttgtctttgc tacttttgtt ccacattctc
2461 tctctttacc tttgccctac cttctgtttg taaggatgat tttaatatta attccaggaa
2521 cacatgtttt tatttctctc gctggaatag gagaaagctt aatatattcc caagctgaat
2581 gttcttgaat tatatctgca tttaccactg tatatgcata tagtgacagt ataacctgtc
2641 tatactacag tagttccagt cctaactttc tgaattatct ttaaagatct tctctaacaa
2701 gctatgggaa tttggcttca tactctttct ttgcaacagc agtgttctgg gtgataatct
2761 tgaattgata cctgttcctt tttctgggtt ttgttggtt tttgaaaaat tgtctttcct
2821 tatcattggg gggaggcttg gtagcaaatg aacatttttt ggaaaagagg acagaaaaat
2881 tgaactacag cttgagaacg tattcttttt ttctactttt gttattgcaa attgaggaat
2941 cacttttaac tgttttaggt gtgtgtgtcc agagtgcaca aggattatgt ttttggattg
3001 tcaaagagga tgcttagtct taaaataaaa ataaatttaa aaatcatctt ataatt

```

FIGURE 61

Figure 62

Name: BIR2Gen Bank Accession No. L49432SEQ ID No.63

```
1 gggcagcagg tttaaaaagg aggaaaacga cttcttctag attttttttt cagtttcttc
61 tataaatcaa aactacctcc cttagagaaag gctagtccct tttcttcccc attcatttca
121 ttatgaacat agtagaaaac agcatattct tatcaaattt gatgaaaagc gccaacacgt
181 ttgaactgaa atacgacttg tcatgtgaac tgtaccgaat gtctacgtat tccacttttc
241 ctgctggggg ccctgtctca gaaaggagtc ttgctcgcgc tggtttctat tacactgggtg
301 tgaatgacaa ggtcaaatgc ttctgttggt gcctgatgct ggataactgg aaaagaggag
361 acagtcctac tgaaaagcat aaaaagttgt atcctagctg cagattcggt cagagtctaa
421 attccgttaa caacttggaa gctacctctc agcctacttt tccttcttca gtaacaaatt
481 ccacacactc attacttccg ggtacagaaa acagtggata tttccgtggc tcttattcaa
541 actctccatc aaatcctgta aactccagag caaatcaaga tttttctgcc ttgatgagaa
601 gttcctacca ctgtgcaatg aataacgaaa atgccagatt acttactttt cagacatggc
661 cattgacttt tctgtcgcca acagatctgg caaaagcagg cttttactac ataggacctg
721 gagacagagt ggcttgcttt gcctgtgggt gaaaattgag caattgggaa ccgaaggata
781 atgctatgtc agaacacctg agacattttc ccaaagccc atttatagaa aatcagcttc
841 aagacacttc aagatacaca gtttctaata tgagcatgca gacacatgca gcccgttcta
901 aaacattctt taactggccc tctagtgttc tagttaatcc tgagcagctt gcaagtgcgg
961 gtttttatta tgtgggtaac agtgatgatg tcaaatgctt ttgctgtgat ggtggactca
1021 ggtgttggga atctggagat gatccatggg ttcaacatgc caagtgggtt ccaaggtgtg
1081 agtacttgat aagaattaaa ggacaggagt tcatccgtca agttcaagcc agttaccctc
1141 atctacttga acagctgcta tccacatcag acagcccagg agatgaaaat gcagagtcac
1201 caattatcca ttttgaacct ggagaagacc attcagaaga tgcaatcatg atgaatactc
1261 ctgtgattaa tgctgccgtg gaaatgggct ttagtagaag cctggtaaaa cagacagttc
1321 agagaaaaat cctagcaact ggagagaatt atagactagt caatgatctt gtgttagact
1381 tactcaatgc agaagatgaa ataagggaag aggagagaga aagagcaact gagggaaaaag
1441 aatcaaatga tttattatta atccggaaga atagaatggc actttttcaa catttgactt
1501 gtgtaattcc aatcctggat agtctactaa ctgccggaat tattaatgaa caagaacatg
1561 atgttattaa acagaagaca cagacgtctt tacaagcaag agaactgatt gatacgattt
1621 tagtaaaagg aaatattgca gccactgtat tcagaaactc tctgcaagaa gctgaagctg
1681 tgttatatga gcatttatatt gtgcaacagg acataaaaata tattcccaca gaagatgttt
1741 cagatctacc agtggagaaa caattgcgga gactacaaga agaaagaaca tgtaaagtgt
1801 gtatggacaa agaagtgtcc atagtgttta ttcttgtgg tcatctagta gtatgcaaag
1861 atttgtctcc ttctttaaga aagtgtccta tttgtaggag tacaatcaag ggtacagttc
1921 gtacatttct tcatgaaga agaaccaaaa catcatctaa actttagaat taatttatta
1981 aatgtattat aactttaact ttcattccta tttggtttcc ttaaaatttt tattttatta
2041 caactcaaca aacattgttt tgtgtaacat atttaataata tgtatctaaa ccatatgaac
2101 atatatTTTT tagaaactaa gagaatgata ggcttttgtt cttatgaacg aaaaagaggt
2161 agcactacaa acacaatatt caatcaaat ttcagcatta ttgaaattgt aagtgaagta
2221 aaacttaaga tatttgagtt aacctttaag aattttaaat attttggcat tgtactaata
2281 ccgggaacat gaagccaggt gtggtggtat gtgcctgtag tcccaggctg aggcaagaga
2341 attacttgag cccaggagtt tgaatccatc ctgggcagca tactgagacc ctgcctttaa
2401 aaacaaacag aacaaaaaca aaacaccagg gacacatttc tctgtctttt ttgatcagtg
2461 tcctatacat cgaaggtgtg catatatgtt gaatgacatt ttagggacat ggtgttttta
2521 taaagaattc tgtgagaaaa aatttaataa aaccccccaa att
```

FIGURE 62

Figure 63

Name: BIR3

Gen Bank Accession No. L49431

SEQ ID No.

64

```

1 tctaagtagt atcttggaaa ttcagagaga tactcatcct acctgaatat aaactgagat
61 aaatccagta aagaaagtgt agtaaattct acataagagt ctatcattga tttcttttgg
121 tggtaaaaaat cttagtccat gtgaagaaat ttcattgtgaa tgttttagct atcaaacagc
181 actgtcacct actcatgcac aaaactgcct cccaaagact tttcccaggt cctcgtatc
241 aaaacattaa gagtataatg gaagatagca cgatcttgtc agattggaca aacagcaaca
301 aacaaaaaat gaagtatgac ttttcctgtg aactctacag aatgtctaca tattcaactt
361 tccccgccgg ggtgccgtgc tcagaaagga gtcttgctcg tgctggtttt tattatactg
421 gtgtgaatga caaggtcaaa tgcttctgtt gtggcctgat gctggataac tggaaactag
481 gagacagtcc tattcaaaag cataaacagc tatatcctag ctgtagcttt attcagaatc
541 tggtttcagc tagtctggga tccacctcta agaatacgtc tccaatgaga aacagttttg
601 cacattcatt atctcccacc ttggaacata gtagcttggt cagtggttct tactccagcc
661 tttctccaaa cctcttaat tctagagcag ttgaagacat ctcttcacg aggactaacc
721 cctacagtta tgcaatgagt actgaagaag ccagatttct tacctaccat atgtggccat
781 taactttttt gtcaccatca gaattggcaa gagctggttt ttattatata ggacctggag
841 atagggtagc ctgctttgcc tgtggtgagg agctcagtaa ctgggaacca aaggatgatg
901 ctatgtcaga acaccggagg cattttccca actgtccatt tttggaaaat tctctagaaa
961 ctctgaggtt tagcatttca aatctgagca tgcagacaca tgcagctcga atgagaacat
1021 ttatgtactg gccatctagt gtccagttc agcctgagca gcttgcaagt gctggttttt
1081 attatgtggg tcgcaatgat gatgtcaaat gcttttggtg tgatgggtggc ttgaggtggt
1141 gggaatctgg agatgatcca tgggtagaac atgccaaagt gtttccaagg tgtgagttct
1201 tgatacgaat gaaaggccaa gagtttggtg atgagattca aggtagatat cctcatcttc
1261 ttgaacagct gttgtcaact tcagatacca ctggagaaga aaatgctgac ccaccaatta
1321 ttcatTTTTG acctggagaa agttcttcag aagatgctgt catgatgaat acacctgtgg
1381 ttaaactctgc cttggaaatg ggctttaata gagacctggt gaaacaaaca gttcaaagta
1441 aaatcctgac aactggagag aactataaaa cagttaatga tattgtgtca gcacttctaa
1501 atgctgaaga tgaaaaaaga gaggaggaga aggaaaaaca agctgaagaa atggcatcag
1561 atgatttgct attaattcgg aagaacagaa tggctctctt tcaacaattg acatgtgtgc
1621 ttctatcctt ggataatctt ttaaaggcca atgtaattaa taaacaggaa catgatatta
1681 ttaaacaaaa aacacagata cctttacaag cgagagaact gattgatacc attttggtta
1741 aaggaaatgc tgcggccaac atcttcaaaa actgtctaaa agaaattgac tctacattgt
1801 ataagaactt atttgtggat aagaatatga agtatattcc aacagaagat gtttcagggtc
1861 tgtcactgga agaacaattg aggagggtgc aagaagaacg aacttgtaaa gtgtgtatgg
1921 acaagaagt ttctgttgta ttatttcctt gtggtcatct ggtagtatgc caggaatgtg
1981 ccccttctct aagaaaatgc cctatttgca ggggtataat caagggtact gttcgtacat
2041 ttctctctta aagaaaaata gtctatatat taacctgcat aaaaaggctt ttaaaatatt
2101 gttgaacact tgaagccatc taaagtaaaa agggaattat gagtttttca attagtaaca
2161 ttcatgttct agtctgcttt ggtactaata atcttgtttc tgaaaagatg gtatcatata
2221 tttaatctta atctgtttat ttacaaggga agatttatgt ttggtgaact atattagtat
2281 gtatgtgtac ctaaggaggat agtgtcactg cttgttatgc atcatttcag gatttactgg
2341 atttgttggt ctttcagaaa gctttgaata ctaaattata gtgtagaaaa gaactggaaa
2401 ccaggaactc tggagttcat cagagttatg gtgccgaatt gtctttgggt cttttcactt
2461 gtgtttttaa ataaggattt ttctcttatt tctccccta gtttgtgaga aacatctcaa
2521 taaagtgcct t

```

FIGURE 63

Figure 64

Name: CBLGen Bank Accession No. X57110SEQ ID No.65

```
1 gaattccggg cccgcatagc cggcgccggc ggccggcgcg gcggcgccgg cggccgggag
61 agggccctcc ttccagccct gcttctctcc ctgctcgca gtcgagccga gccggcgga
121 cgccttgggc tccgaccctg cccaggccat ggccggcaac gtgaagaaga gctctggggc
181 cggggcgccg acgggctccg ggggctcggg ttcgggtggc ctgattgggc tcatgaagga
241 cgccttccag ccgcaccacc accaccacca ccacctcagc cccaccgccg cggggacggt
301 ggacaagaag atggtggaga agtgcctgga gctcatggac aaggtggtgc ggttgtgtca
361 gaacccaaag ctggcgctaa agaatagccc accttatatc ttagacctgc taccagatac
421 ctaccagcat ctccgtacta tcttgtcaag atatgagggg aagatggaga cacttggaga
481 aaatgagtat tttaggggtg ttatggagaa tttgatgaag aaaactaagc aaaccataag
541 cctcttcaag gagggaaaag aaagaatgta tgaggagaat tctcagccta ggcgaaacct
601 aaccaaactg tccctcatct tcagccacat gctggcagaa ctaaaaggaa tctttccaag
661 tggactcttt caggagaca catttcggat tactaaagca gatgctgcgg aattttggag
721 aaaagctttt ggggaaaaga caatagtccc ttggaagagc tttcgacagc ctctacatga
781 agtgcattcc atcagttctg ggctggaggc catggctctg aaatccacta ttgatctgac
841 ctgcaatgat tataatttcg tttttgaatt tgacatcttt acccgactct ttcagccctg
901 gtcctctttg ctcaggaatt ggaacagcct tgctgtaact catcctggct acatggcttt
961 tttgacgtat gacgaagtga aagctcggct ccagaaattc attcacaac ctggcagtta
1021 tatcttccgg ctgagctgta ctgctctggg tcagtgggct attgggtatg ttactgctga
1081 tgggaacatt ctccagacaa tccctcaca taaacctctc ttccaagcac tgattgatgg
1141 cttcaggga ggcttctatt tgtttcctga tggacgaaat cagaatcctg atctgactgg
1201 cttatgtgaa ccaactcccc aagaccatat caaagtgacc caggaacaat atgaattata
1261 ctgtgagatg ggctccacat tccaactatg taaaatatgt gctgaaaatg ataaggatgt
1321 aaagattgag cctgtgggac acctcatgtg cacatcctgt cttacatcct gccaggaatc
1381 agaaggctcag ggctgtcctt tctgccgatg tgaaattaaa ggtactgaac ccatcgtggt
1441 agatccgttt gatcctagag ggagtggcag cctgttgagg caaggagcag agggagctcc
1501 ctccccaat tatgatgatg atgatgatga acgagctgat gatactctct tcatgatgaa
1561 ggaattggct ggtgccaaag tggaacggcc gccttctcca ttctccatgg cccacaagc
1621 ttcccttccc cgggtgccac cagacttga ccttctgccg cagcgagtat gtgttccctc
1681 aagtgtctct gctcttggaa ctgcttctaa ggctgtctct ggctcccttc ataaagacaa
1741 accattgcca gtacctcca cacttcgaga tcttcacca ccaccgctc cagaccggcc
1801 atattctgtt ggagcagaat cccgacctca aagacgccc ttgccttgta caccaggcga
1861 ctgtccctcc agagacaaac tgccccctgt cccctctagc cgccttggag actcatggct
1921 gccccggcca atcccaaaag taccagtatc tgccccaagt tccagtgatc cctggacagg
1981 aagagaatta accaaccggc actcacttcc attttcattg ccctcacaaa tggagcccag
2041 accagatgtg cctaggctcg gaagcacgtt cagtctggat acctccatga gtatgaatag
2101 cagcccatga gtaggtccag agtgtgacca cccaaaatc aaaccttct catctgccaa
2161 tgccatttat tctctggctg ccagacctct tctgtgcca aaactgccac ctggggagca
2221 atgtgagggt gaagaggaca cagagtacat gactccctct tccaggctc tacggccttt
2281 ggatacatcc cagagttcac gagcatgtga ttgcgaccag cagattgata gctgtacgta
2341 tgaagcaatg tataatattc agtcccagc gccatctatc accgagagca gcacctttgg
2401 tgaaggaat ttggccgag cccatgccaa cactggctcc gaggagtcag aaaaatgagga
2461 tgatgggtat gatgtcccaa agccacctgt gcggcgctg ctggcccgcc gaactctctc
2521 agatatctct aatgccagct cctcctttgg ctggttgtct ctggatggtg atcctacaac
2581 aaatgtcact gaaggttccc aagttcccga gaggcctcca aaaccattcc cgcggagaat
2641 caactctgaa cggaaagctg gcagctgtca gcaaggtagt ggtcctgccc cctctgctgc
2701 caccgctca cctcagctct ccagtgaat cgagaacctc atgagtcagg ggtactccta
2761 ccaggacatc cagaaagctt tggctattgc ccagaacaac atcgagatgg ccaaaaacat
2821 cctccgggaa tttgtttcca tttcttctcc tgcccatgta gctacctagc acaccatctc
2881 cctgctgcag gtttagagga ccagtgaatt gggagtatt actcaagtgg cactagaag
2941 ggcaggagtt cctttggtga cttcacagt aagtcttgcc ctctctgtgg gatattcacat
3001 cagtggttcc aagatttcaa agtggtgaaa tgaaaatgga gcagctagta tgttttatta
3061 ttttatgggt cttgagtga tttgaagtg
```

FIGURE 64

Figure 65

Name: PMLGen Bank Accession No. M79462SEQ ID No.66

1 ctaagctggg gtccatggag cctgcacccg cccgatctcc gaggccccag caggaccccc
61 cccggcccca ggagcccacc atgcctcccc ccgagacccc ctctgaaggc cgccagccca
121 gccccagccc cagccctaca gagcgagccc ccgcttcgga ggaggagttc cagtttctgc
181 gctgccagca atgccaggcg gaagccaagt gcccgaagct gctgccttgt ctgcacacgc
241 tgtgctcagg atgcctggag gcgtcgggca tgcagtgcgc catctgccag gcgcctggc
301 ccctaggtgc agacacaccc gccctggata acgtcttttt cgagagtctg cagcggcgcc
361 tgtcgggtga ccggcagatt gtggatgcgc aggctgtgtg caccgcgtgc aaagagtcgg
421 ccgacttctg gtgctttgag tgcgagcagc tccctgtgcg caagtgcctc gaggcacacc
481 agtgggttct caagcacgag gcccggcccc tagcagagct gcgcaaccag tcggtgcgtg
541 agttcctgga cggcacccgc aagaccaaca acatcttctg ctccaacccc aaccaccgca
601 cccctacgct gaccagcatc tactgccgag gatgttccaa gccgctgtgc tgctcgtgcg
661 cgctccttga cagcagccac agtgagctca agtgcgacat cagcgcagag atccagcagc
721 gacaggagga gctggacgcc atgacgcagg cgctgcagga gcaggatagt gcctttggcg
781 cggttcacgc gcagatgcac gcggccgtcg gccagctggg ccgcgcgcgt gccgagaccg
841 aggagctgat ccgcgagcgc gtgcgccagg tggtagctca cgtgcgggct caggagcgcg
901 agctgctgga ggctgtggac gcgcggtacc agcgcgacta cgaggagatg gccagtccgc
961 tgggcgcgct ggatgctgtg ctgcagcgca tccgcacggg cagcgcgctg gtgcagagga
1021 tgaagtgcta cgcttcggac caggaggtgc tggacatgca cggtttctctg cgccaggcgc
1081 tctgccgcct gcgccaggag gagccccaga gcctgcaagc tgccgtgcgc accgatggct
1141 tcgacgagtt caaggtgcgc ctgcaggacc tcagctcttg catcacccag gggaaaagatg
1201 cagctgtatc caagaaagcc agcccagagg ctgccagcac tcccagggac cctattgacg
1261 ttgacctggc cgaggaggca gagagagtga agggccaggc tcaggccctg gggctggtg
1321 aagcccagcc tatggctgtg gtacagtacg tggccggggc acaccccgctg ccagtgtacg
1381 ccttctccat caaaggccct tccatggag aggtgtctc caatacaacg acagcccaga
1441 agaggaagtg cagccagacc cagtgcacca ggaaggctcat caagatggag tctgaggagg
1501 ggaaggaggc aaggttggct cggagctccc cggagcagcc caggcccagc acctccaagg
1561 cagtctcacc accccacctg gatggaccgc ctagccccag gagccccgtc ataggaagtg
1621 aggtcttctc gcccaacagc aaccacgtgg ccagtggcgc cggggaggca gaggaacgcg
1681 ttgtggtgat cagcagctcg gaagactcag atgccgaaaa ctctctctcc cgagagctgg
1741 atgacagcag cagtgagtc agtgacctcc agctgggaagg cccagcacc ctgagggtcc
1801 tggacgagaa ccttgctgac cccaagcag aagacagacc tctggttttc tttgacctca
1861 agattgacaa tgaaacccag aagattagcc agctggctgc ggtgaaccgg gaaagcaagt
1921 tccgcgtggt catccagcct gaagccctct tcagcatcta ctccaaggcc gtgtccctgg
1981 aggtggggct gcagcacttc ctgagcttct tgagctccat gcgccgccc atcttgacct
2041 gctacagagc gtgggggcct ggcctcccaa acttcttccg ggccctggag gacattgaaca
2101 ggctgtggga attccaggag gccatctcgg gcttctggc tgccctgcct ctcatccggg
2161 agcgtgtgcc cggggccagc agcttcaaac tcaagaacct ggcccagacc tacctggcga
2221 gaaacatgag cgagcgcagc gccatggctg ccgtgctggc catgcgtgac ctgtgccgcc
2281 tccctgaggt cccccgggc cccagctgg cccagcatgt ctaccccttc agtagcctgc
2341 agtgccttgc ctccctgcag ccctgggtgc aggcagctgt gctgccccgg gctgaggccc
2401 gcctcctggc cctacacaac gtgagcttca tggagctgct gagtgcacac cgccgtgacc
2461 ggcagggggg cctgaagaag tacagccgct atctaagcct gcagaccacc acgttgcccc
2521 ctgcccagcc tgctttcaac ctgcaggctc tgggcaacct ctttgaaggc ctgttggagg
2581 gtccggcgct ggcacgggca gaaggagtct ccacccccact tgcctggcgt ggcttggcag
2641 agagggcctc ccagcagagc tgagaggagg gggtgaccag cttggagtct ctggtgggca
2701 gagagggatg gggctccctga gccaggcccc acccatcaca gcattcccag gtcttggtac
2761 ccagccctca gttgtcattt ggttcagaat cagttccctt tctctgggac caaatttccc
2821 ttctctaata atcctacaga gaaggttcca aagctgagca cccatcaccc taggtgtgca
2881 ccagactcct attagccct ccttcaggc gctagaacca gggaggtcct gagggtgaa
2941 ggcagacact ctgggctctc taggtggccc tggctcttgc ccatccatcc cacctgccac
3001 taccttgggt gtccttacag ctctgccctg accccagccc ctgcccctgg acacctgctg

```
3061 acagctccca cacagaacag tctgaggtga tgctggctac agccctggca gcccattggca
3121 actcaaagtg ccttccaaac caaactgccc ctcatgaggt taggggtccc caccctccac
3181 ctgcccagag gccaaactctg ttcccttctc ctttcatccc agaggggcct catcagcaaa
3241 gacagaaact gatgttccat gcatgtccct gcttccaaag cctgaccttc caaagcttgc
3301 ttcccttctc ctggctgcac agacacctct gcttggccac agctttgctc tttggtcctc
3361 aggccaccag acccctcaca gcatggccct tggtctttcc tgcactgccc tcaccctcag
3421 ccgtcccaag gagtgccag cactactcag catgtagtcc aggacctgcg gcatcagcat
3481 cccctgggag cttcttagaa atgcagacct gtgggctcta ccacaggcct ctggaatcag
3541 aatcttcagg gtggggccca gcagtctgtg ttttaacagg ttctcccggt gatgctggtg
3601 catgctcaag tttgagaacc gctgctctca tagactgctc ctccaagggg aagcagtgtg
3661 gaacagcaga gaaagtccgt ccctatctct gactgtgcag tgagtgtggc ctaggaacag
3721 gtcccttctc aagctctagt gtcccatct gttaaagggt tgattggcct ccacggtcag
3781 agcggccatc tggctgtgca tctgtcattt ttaggaatgg aaagcaggcc tctgaggcag
3841 tggacaggaa gacttctcat ccttgaattc tagctccatt ccaaactggt agtccaacct
3901 tgtggtcaca tcagaggtca acagcagaac agaggcagggt cagggttgtc cgctctgatt
3961 agcagatcga cagccccctc ctccctctcc ctcttccca tccctactca ttagctccct
4021 cgtcctggga tgcttocaag cagcccagggt ctctagcctc catggctgat gacctcagct
4081 ttcacactgg tgaagattgt gtaccatac tgcagcccca tcccatgggc ccctgaagac
4141 ccactgagga attccgtagg tcttgttcca cgaccggagt gctggctctc acagtgaatt
4201 ttgatgcatt taaaataaga ttctgatgcc agactgttaa aacaggcgct ggttctgaag
4261 caccatggaa acagattgat gcagatggaa ggaggagggt agactggggt cactgatttc
4321 cagccccacc atgtgtgctg ttttcagatg tgattgagggt gtgttctgcc ctgcctccac
4381 tgtcacagcc ttttaataaag atgtgaatct tgaaaaaaaa aaaaaaaaaa aaaaaaaaaa
4441 aaaaaaaaaa aaaaaaaaaa aaa
```

FIGURE 65

Figure 66

Name: RAG1Gen Bank Accession No. M29474SEQ ID No. 67

```
1 gagagcagag aacacacttt gccttctctt tggatttgag taatatcaac caaattgcag
61 acatctcaac actttggcca ggcagcctgc tgagcaaggt acctcagcca gcatggcagc
121 ctctttccca cccaccttgg gactcagttc tgcccagat gaaattcagc acccacatat
181 taaattttca gaatggaaat ttaagctgtt ccgggtgaga tcctttgaaa agacacctga
241 agaagctcaa aaggaaaaga aggattcctt tgaggggaaa ccctctctgg agcaatctcc
301 agcagtcctg gacaaggetg atggtcagaa gccagtccca actcagccat tgttaaaagc
361 ccaccetaag ttttcaaaga aatttcacga caacgagaaa gcaagaggca aagcgatcca
421 tcaagccaac cttcgacatc tctgccgat ctgtgggaat tcttttagag ctgatgagca
481 caacaggaga tatccagtcc atggctcctg ggatggtaaa accctaggcc ttttacgaaa
541 gaaggaaaag agagctactt cctggccgga cctcattgcc aagggtttcc ggatcgatgt
601 gaaggcagat gttgactcga tccacccac tgagttctgc cataactgct ggagcatcat
661 gcacaggaag tttagcagtg ccccatgtga ggtttacttc ccgaggaacg tgaccatgga
721 gtggcaccac cacacacccat cctgtgacat ctgcaacact gcccgtcggg gactcaagag
781 gaagagtctt cagccaaact tgcagctcag caaaaaactc aaaactgtgc ttgaccaagc
841 aagacaagcc cgtcagcgca agagaagagc tcaggcaagg atcagcagca aggatgtcat
901 gaagaagatc gccaaactgca gtaagataca tcttagtacc aagctccttg cagtggactt
961 cccagagcac tttgtgaaat ccattccttg ccagatctgt gaacacattc tggctgacct
1021 tgtggagacc aactgtaagc atgtcttttg ccgggtctgc attctcagat gcctcaaagt
1081 catgggcagc tattgtccct cttgccgata tccatgcttc cctactgacc tggagagtcc
1141 agtgaagtcc tttctgagcg tcttgaattc cctgatgggtg aaatgtccag caaaagagtg
1201 caatgaggag gtcagtttgg aaaaatataa tcaccacatc tcaagtcaca aggaatcaaa
1261 agagattttt gtgcacatta ataaaggggg ccggcccgcc caacatcttc tgtcgctgac
1321 tcggagagct cagaagcacc ggctgaggga gctcaagctg caagtcaaag cctttgctga
1381 caaagaagaa ggtggagatg tgaagtccgt gtgcatgacc ttgttcctgc tggctctgag
1441 ggcgaggaat gagcacaggc aagctgatga gctggaggcc atcatgcagg gaaagggctc
1501 tggcctgcag ccagctgttt gcttggccat ccgtgtcaac accttcctca gctgcagtca
1561 gtaccacaag atgtacagga ctgtgaaagc catcacaggg agacagattt ttcagccttt
1621 gcatgccctt cggaatgctg agaaggtact tctgccaggc taccaccact ttgagtggca
1681 gccacctctg aagaatgtgt cttccagcac tgatgttggc attattgatg ggtgtctggtg
1741 actatcatcc tctgtggatg attaccagat ggacaccatt gcaaagaggt tccgctatga
1801 ttcagctttg gtgtctgctt tgatggacat ggaagaagac atcttggaag gcatgagatc
1861 ccaagacctt gatgattacc tgaatggccc cttcactgtg gtggtgaagg agtcttgtga
1921 tggaaatggga gacgtgagtg agaagcatgg gagtgggctt gtagttccag aaaaggcagt
1981 ccgttttttca ttcacaatca tgaaaattac tattgccac agctctcaga atgtgaaagt
2041 atttgaagaa gccaaaccta actctgaact gtgttgcaag ccattgtgcc ttatgctggc
2101 agatgagtct gaccacgaga cgctgactgc catcctgagt cctctcattg ctgagaggga
2161 ggccatgaag agcagtgaat taatgcttga gctgggaggc attctccgga ctttcaagtt
2221 catcttcagg ggcaccgct atgatgaaaa acttgtgcgg gaagtggaag gcctcgaggc
2281 ttctggctca gtctacattt gtactctttg tgatgccacc cgtctggaag cctctcaaaa
2341 tcttgtcttc cactctataa ccagaagcca tgctgagaac ctggaacgtt atgaggtctg
2401 gcgttccaac ccttaccatg agtctgtgga agaactgcgg gatcgggtga aaggggtctc
2461 agctaaacct ttcattgaga cagtcccttc catagatgca ctccactgtg acattggcaa
2521 tgcagctgag ttttacaaga tcttccagct agagataggg gaagtgtata agaattccaa
2581 tgcttccaaa gaggaaggga aaaggtggca ggcacactg gacaagcatc tccggaagaa
2641 gatgaacctc aaaccaatca tgaggatgaa tggcaacttt gccaggaagc tcatgaccaa
2701 agagactgtg gatgcagttt gtgagttaat tccttccgag gagaggcacg aggtctgag
2761 ggagctgatg gatctttacc tgaagatgaa accagtatgg cgatcatcat gccctgctaa
2821 agagtgccca gaatccctct gccagtacag tttcaattca cagcgttttg ctgagctcct
2881 ttctacgaag ttcaagtata ggtatgaggg aaaaatcacc aattattttc acaaaacctt
2941 ggcccatgtt cctgaaatta ttgagaggga tggctccatt ggggcatggg caagtggggg
3001 aaatgagtct ggtaacaaac tgtttaggcg cttccggaaa atgaatgcca ggcagtccaa
3061 atgctatgag atggaagatg tcttgaaaca ccactgggtt tacacctcca aatacctcca
3121 gaagtttatg aatgctcata atgcattaaa aacctctggg tttaccatga accctcaggc
3181 aagcttaggg gaccattag gcatagagga ctctctggaa agccaagatt caatggaatt
```

```

3241 ttaagtaggg caaccactta tgagttgggt tttgcaattg agtttccctc tgggttgcat
3301 tgagggtctt ccttagcacc ctttactgct gtgtatgggg cttcaccatc caagagggtg
3361 taggttggag taagatgcta cagatgctct caagtcagga atagaaactg atgagctgat
3421 tgcttgaggc ttttagtgag ttccgaaaag caacaggaaa aatcagttat ctgaaagctc
3481 agtaactcag aacaggagta actgcagggg accagagatg agcaaagatc tgggtgtggt
3541 ggggagctgt catgtaaatc aaagccaagg ttgtcaaaga acagccagtg aggccagaaa
3601 ttggtcttgt ggttttcatt tttttccccc ttgattgatt atattttgta ttgagatatg
3661 ataagtgcct tctatttcat ttttgaataa ttcttcattt ttataatttt acatatcttg
3721 gcttgctata taagattcaa aagagctttt taaatttttc taataatata ttacatttgt
3781 acagcatgat gacctttaca aagtgcctct aatgcattta cccattcggt atataaatat
3841 gttacatcag gacaactttg agaaaatcag tcctttttta tgtttaaatt atgtatctat
3901 tgtaaccttc agagtttagg aggtcatctg ctgtcatgga tttttcaata atgaatttag
3961 aatacacctg ttagctacag ttagttatta aatcttctga taatatatgt ttacttagct
4021 atcagaagcc aagtatgatt ctttattttt actttttcat ttcaagaaat ttagagtttc
4081 caaattttaga gcttctgcat acagtcttaa agccacagag gcttgtaaaa atataggtta
4141 gcttgatgtc taaaaatata tttcatgtct tactgaaaca ttttgccaga ctttctccaa
4201 atgaaacctg aatcaatttt tctaaatcta ggtttcatag agtcctctcc tctgcaatgt
4261 gttattcttt ctataatgat cagtttactt tcagtggatt cagaattgtg tagcaggata
4321 acctgtatt tttccatccg ctaagtttag atggagtcca aacgcagtac agcagaagag
4381 ttaacattta cacagtgcct ttaccactg tggaatgttt tcacactcat ttttccttac
4441 aacaattctg aggagtaggt gttgttatta tctccatttg atgggggttt aatgatttgc
4501 tcaaagtcat ttaggggtaa taaatacttg gcttggaat ttaacacagt ctttttgtct
4561 ccaaagccct tcttctttcc accacaaatt aatcactatg tttataagggt agtatcagaa
4621 tttttttagg attcacaact aatcactata gcacatgacc ttgggattac atttttatgg
4681 ggcaggggta agcggctttt aaatcatttg tgtgctctgg ctcttttgat agaagaaagc
4741 aacacaaaag ctccaaaggg cccctaaccc ctcttggtgg tccagtattt tggaaactat
4801 gatctgcatc cttaggaatc tgggatttgc cagttgctgg caatgtagag caggcatgga
4861 attttatatg ctagtgagtc ataatgatat gttagtgtta attagttttt cttcctttga
4921 ttttattggc cataattgct actcttcata cacagtatat caaagagctt gataatttag
4981 ttgtcaaaag tgcacggcg acattatctt taattgtatg tatttggtgc ttcttcaggg
5041 attgaactca gtatctttca ttaaaaaaca cagcagtttt ccttgctttt tatatgcaga
5101 atatcaaagt catttctaata ttagtgttca aaaacatata catattttta cattagtttt
5161 tttgaaaact cttggttttg tttttttgga aatgagtggg ccactaagcc acactttccc
5221 ttcaccttgc ttaatccttc cagcatgtct ctgcactaat aaacagctaa attcacataa
5281 tcatcctatt tactgaagca tggcatgctt ggtttataga ttttttacct atttctactc
5341 tttttctcta ttggtggcac tgtaaatact ttccagtatt aaattatcct tttotaacac
5401 tgtaggaact attttgaatg catgtgacta agagcatgat ttatagcaca acctttccaa
5461 taatccctta atcagatcac attttgataa accctgggaa catctggctg caggaatttc
5521 aatatgtaga aacgtgcct atggtttttt gcccttactg ttgagactgc aatatcctag
5581 accctagttt tatactagag ttttattttt agcaatgcct attgcaagtg caattatata
5641 ctccagggaa attcaccaca ctgaatcgag catttgtgtg tgtatgtgtg aagtatatct
5701 gggacttcag aagtgcattg tatttttctc ctgtgaaacc tgaatctaca agttttctgc
5761 caagccactc aggtgcattg cagggaccag tgataatggc tgatgaaaat tgatgattgg
5821 tcagtgaggt caaaaggagc cttgggatta ataaacatgc actgagaagc aagaggagga
5881 gaaaaagatg tctttttctt ccaggtgaac tggaatttag ttttgctca gatttttttc
5941 ccacaagata cagaagaaga taaagatttt tttggttgag agtgtgggtc ttgcattaca
6001 tcaaacagag ttcaaattcc acacagataa gaggcaggat atataagcgc cagtggtagt
6061 tgggaggaat aaaccattat ttggatgcag gtggtttttg attgcaaata tgtgtgtgtc
6121 ttcagtgatt gtatgacaga tgatgtattc ttttgatgtt aaaagatttt aagtaagagt
6181 agatacattg tacccatttt acattttctt attttaacta cagtaatcta cataaatata
6241 cctcagaaat catttttggg gattattttt tgtttttag aattgcactt cagtttattt
6301 tcttacaaat aaccttacet tttgtttaat ggcttccaag agcctttttt tttttgtatt
6361 tcagagaaaa ttcaggtacc aggatgcaat ggatttattt gattcagggg acctgtattt
6421 ccatgtcaaa tgttttcaaa taaaatgaaa tatgagtttc aatactttt atattttaat
6481 atttccttaa tattatgggt attgtccgcc attttgttgt atattgtaaa taaagtttag
6541 attgt

```

FIGURE 66

Figure 67

Name: RAPSYNGen Bank Accession No. Z33905SEQID No. 68

```
1  cccaactggc agcgacacgt agggacgggc tgaaccagct ttcttcccag ggtggcgccct
61  gctctccatc caggcccccatt tccggctccc acccgactgc tgcttttgtt cccacgtttc
121  ggggggcagc tggcactgtg attcctgccc catgagtgcc tagtggccag gagccaccag
181  ggatcacccc acgtgggctt ggggaggatg gggcaggacc agaccaagca gcagatcgag
241  aaggggctcc agctgtacca gtccaaccag acagagaagg cattgcagggt gtggacaaag
301  gtgctggaga agagctcgga cctcatgggg cgcttccgcg tgctgggctg cctggtcaca
361  gccactcgg agatgggccc ctacaaggag atgctgaagt tcgccgtggt ccagatcgac
421  acggcccggg agctggagga tgccgacttc ctccctggaga gctacctgaa cctggcacgc
481  agcaacgaga agctgtgcga gtttcacaag accatctcct actgcaagac ctgccttggc
541  gtgccaggta ccaggggcagg tgcccagctc ggaggccagg tcagcctgag catgggcaat
601  gccttcttgg gcctcagcgt cttccagaag gccctggaga gcttcgagaa ggccctgcgc
661  tacgcgcaca acaatgatga caccatgctc gagtgccgcg tgtgctgcag cctgggcagc
721  ttctatgccc aggtcaagga ctacgagaaa gccctgttct tcccctgcaa ggcggcagag
781  cttgtcaaca actatggcaa aggtctggagc ctgaagtacc aggccatgag ccagtaccac
841  atggccgtgg cctatgcctt gctgggcccgc ctgggcagtg ccatggagtg ttgtgaggag
901  tctatgaaga tcgccctgca gcacggcgac cggccactgc aggcgctctg cctgctctgc
961  ttcgctgaca tccaccggag ccgtggggac ctggagacag ccttccccag gtacgactcc
1021 gccatgagca tcatgaccga gatcggaaac cgcctggggc aggtgcaggc gctgctgggt
1081 gtggccaagt gctgggtggc caggaaggcg ctggacaagg ctctggatgc catcgagaga
1141 gcccaggatc tggccgagga ggtggggaac aagctgagcc agctcaagct gcaactgtctg
1201 agcgagagca ttaccgcag caaagggtcg cagcgggaac tgccgcgcga cgttgtgagg
1261 ttccacgagt gcgtggagga gacggagctc tactgcgcgc tgtgcggcga gtccataggc
1321 gagaagaaca gccggctgca ggccctaccc tgctcccaca tcttccacct caggtgcctg
1381 cagaacaacg ggacccggag ctgtcccaac tgccgcgcgt cctccatgaa gcctggcttt
1441 gtatgactcc tggcaggagg gcgtgggctt cctcctgggc gactcctgct ctttctccac
1501 tgcaccgagt ggccattact cctgggcagc tgccaggtct gcctcaccat acgcaacgct
1561 tggggccccc agggcgtgct cccctggccc agctcccctg ccctgcctct ttgtactttg
1621 ctctttatag aaaaataaac tgtttgtacc tgg
```

FIGURE 67

Figure 68

Name: TIFI-ALPHA

Gen Bank Accession No. AF009353

SEQ ID No.

69

```

1 atggagggtgg cgggtggagaa ggcggtggcg gcggcgccac ggctcggtcg cgctccgggg
61 ggccctcggg gcggcgggga gaacgaggcc gagagtcggc agggcccggga ctcgagcgcg
121 ggccggcagg cgcccggtct caacctgttg gaccttgcg ccgtgtgcca ccagaacatc
181 cagagccggg cggccaagct gctgccctgc ctgactctt tctgccagcg ctgcctgccc
241 gcgccccagc gctacctcat gctgcccgcg cccatgctgg gctcggccga gaccccgcca
301 cccgtccctg ccccggtctg ccggtcagcc tcgtcgccgt tcgccacca agttggagtc
361 attcgttgcc cagtttgagc ccaagaatgt gcagagagac acatcataga taactttttt
421 gtgaaggaca ctactgaggt tcccagcagt acagtagaaa agtcaaatca ggtatgtaca
481 agctgtgagg acaacgcaga agccaatggg ttttgtgtag agtgtgttga atggctctgc
541 aagacgtgta tcagagctca tcagagggtg aagttcaca aagaccacac tgtcagacag
601 aaagaggaag tatctccaga ggcagttggt gtcaccagcc agcgaccagt gttttgtcct
661 tttcataaaa aggagcagct gaagctgtac tgtgagacat gtgacaaact gacatgtcga
721 gactgtcagt tgtagaaca taaagagcat agataccaat ttatagaaga agcttttcag
781 aatcagaaag tgatcataga tacactaatc accaaactga tggaaaaaac aaaatacata
841 aaattcacag gaaatcagat ccaaacacaga attattgaag taaatcaaaa tcaaaagcag
901 gtggaacagg atattaaagt tgctatatat aactgatgg tagaaataaa taaaaagga
961 aaagctctac tgcacagtt agagagcctt gcaaaggacc atcgcatgaa acttatgcaa
1021 caacaacagg aagtgactgg actctctaaa caattggagc atgtcatgca tttttctaaa
1081 tgggcagttt ccagtggcag cagtacagca ttactttata gcaaacgact gattacatac
1141 cggttacggc acctccttgc tgcaaggtgt gatgcatccc cagtgaacca caacaccatc
1201 caatttcact gtgatcctag tttctgggct caaaatatca tcaacttagg ttcttttagta
1261 atcgaggata aagagagcca gccacaaatg cctaagcaga atcctgtcgt ggaacagaat
1321 tcacagccac caagtgggtt atcatcaaac cagttatcca agttcccaac acagatcagc
1381 ctagctcaat tacggctcca gcatatgcag caacagcaac cgcctccacg tttgataaac
1441 tttcagaatc acagcccca acccaatgga ccagttcttc ctctcatcc tcaacaactg
1501 agatatccac caaaccagaa cataccacga caagcaataa agccaaaccc cctacagatg
1561 gctttcttgg ctcaacaagc cataaaacag tggcagatca gcagtggaca gggaaaccca
1621 tcaactacca acagcacatc ctctactcct tccagcccca cgattactag tgcagcagga
1681 tataatggaa aggtttttgg ttcacctata atcgatttga gtcaccaggt gggagggtct
1741 tataatcttc cctctcttcc ggatattgac tgttcaagta ctattatgct ggacaatatt
1801 gtgaggaaag atactaatat agatcatggc cagccaagac caccctcaaa cagaacggtc
1861 cagtcaccaa attcatcagt gccatctcca ggccttgtag gacctgttac tatgactagt
1921 gtacaccccc caatacgttc acctagtgcc tccagcgttg gaagccgagg aagctctggc
1981 tcttcagca aaccagcagg agctgactct acacacaaag tcccagtggt catgctggag
2041 ccaattcgaa taaaacaaga aaacagtgga ccaccggaaa attatgattt cccagttggt
2101 atagtgaagc aagaatcaga tgaagaatct aggcctcaaa atgccaatta tccaagaagc
2161 atactcacct cctgctctt aaatagcagt cagagctcta cttctgagga gactgtgcta
2221 agatcagatg cccctgatag tacaggagat caacctggac ttcaccagga caattcctca
2281 aatggaaagt ctgaatggtt ggatccttcc cagaagtcac ctcttcatgt tggagagaca
2341 aggaaagagg atgaccccaa tgaggactgg tgtgcagttt gtcaaaacgg aggggaactc
2401 ctctgctgtg aaaagtggcc caaagtattc catctttctt gtcattgtgc caccattgaca
2461 aattttccaa gtggagagtg gatttgcact ttctgccgag acttatctaa accagaagtt
2521 gaatatgatt gtgatgctcc cagtcacaa tcagaaaaaa agaaaactga aggccttggt
2581 aagttaacac ctatagataa aaggagtggt gagcgctac ttttatttct ttactgcat
2641 gaaatgagcc tggcttttca agaccctgtt cctctaactg tgcttgatta ttacaaaata
2701 attaaaaatc caatggattt gtcaaccatc aagaaaagac tacaagaaga ttattccatg
2761 tactcaaaac ctgaagattt tgtacgtgat tttagattga tctttcaaaa ctgtgctgaa
2821 ttcaatgagc ctgattcaga agtagccaat gctggataaa aacttgaaaa ttattttgaa
2881 gaacttctaa agaaccctta tccagaaaaa aggtttccca aaccagaatt caggaatgaa
2941 tcagaagata ataaatttag tgatgattca gatgatgact ttgtacagcc ccggaagaaa
3001 cgcctcaaaa gcattgaaga acgccagttg cttaaataa

```

FIGURE 68

Figure 69

Name: TIFI-GAMMAGen Bank Accession No. AF119043SEQ ID No.70

```

1  cctgcggctg gggctggggg cggcggcggc ggcgacgggg cggcggggcg cgcggggcg
61  tccggcgggt tcaaagagga aaacatggcg gaaaacaaag gcggcggcga ggctgagagc
121 ggcggcgggg gcagcggcag cgcgccggta actgccgggg ccgccggggc cgcgcgcgag
181 gaggcggagc cgctctcac cgcggtgctg gtggaggagg aggaggagga aggcggcagg
241 gccggcgctg agggcggcgc ggcggggccc gacgacgggg gggcggccgc ggcctcctcg
301 ggctcggccc aggtctcttc atctcctgcg gcctcagtgg gcaactggagtg tgcggggggc
361 gcagtatcga cgcgggctcc agctccagcc tcggctcccg ctccgggtcc ctccggcagg
421 ccgcctcctg gaccgccagc ctgcctcctg gacacctgcg ccgtgtgtca gcagagcttg
481 cagagccggc gtgaggcgga gcccagctg ctgccctgtc ttcactcctt ctgctgcgc
541 tgccctgccg agccggagcg ccagctcagc gtgccatcc cggggggcag caacggcgac
601 atccagcaag ttggtgtaac acggcgccca gtatgccgc aagaatgcag acagatagac
661 cttgtggata attattttgt gaaagacaca tctgaagctc ctgacagttc tgatgaaaaa
721 tcagaacagg tatgtactag ttgtgaagac aatgcaagtg cagttggctt ttgtgtagaa
781 tgtggagagt ggctatgtaa gacatgtatc gaagcacatc aaagagtaaa atttactaaa
841 gatcacttga tcaggaagaa agaagatgtc tcagagtctg ttggagcatc tggcacaacg
901 cctgttttct gcctgtaca caaacaagaa cagttgaaac tttctgtga aacatgtgat
961 agattgacat gtagagactg tcagctattg gaacacaaag aacataggta tcagtttttg
1021 gaagaagctt ttcaaaatca gaagggtgca attgagaatc tactggcgaa acttcttgag
1081 aagaagaatt atgttcattt tgcagctact caggtgcaga ataggataaa agaagtaaat
1141 gagactaaca aacgtgtaga acaggaaatt aaagtggcca ttttcacct tatcaatgaa
1201 attaataaga aaggaaaatc tctcttaca cagctagaga atgttacaac ggaaagacag
1261 atgaagttac tacagcagca gaatgacatc acaggccttt cccggcaggt gaagcatgtt
1321 atgaacttca caaattgggc aattgcaagt ggcagcagca cagcactact atacagcaag
1381 cgactgatta ctttccagtt gcgtcatatt ttgaaagcac ggtgtgatcc tgcctctgct
1441 gctaattggag caatacgttt ccattgtgat cccaccttct gggcaagaa tgtagtcaat
1501 ttaggtaatc tagtaataga gagtaaacca gtcctgggtt atactcctaa tgtttagtct
1561 gggcaagttc ctccaggac aaaccacatt agtaaaaccc ctggacagat taacttagca
1621 cagcttcgac tccagcacat gcaacaacaa gtatatgcac agaaacatca gcagttgcaa
1681 cagatgagga tgcagcaacc accagcacct gtaccaacta caacaacaac aacacaacag
1741 catcctagac aagcagcccc tcagatgtta caacaacagc ctctcgatt gatcagtggt
1801 caaacaatgc aaagaggcaa catgaactgt ggagcttttc aagcccatca gatgagactg
1861 gctcagaatg ctgccagaat accagggata cccaggcaca gcggccctca atattccatg
1921 atgcagccac acctccaaag acaacactca aaccaggggc atgctggacc ctttcccgta
1981 gtatcggtac acaacaccac aatcaaccca acgagcccta ctacagcaac tatggcaaat
2041 gcaaaccgag gtcccaccag ccatctgttt acagcaatag agctaattcc ctcatgtacc
2101 aatccagaaa accttccatc gctgccagat attccacca tacagttgga agatgctggc
2161 tcaagtagtt tagataatct actaagtaga tacatctcag gcagtcacct accccacag
2221 cctacaagca ccatgaatcc ttctccaggt cctctgccc tttctccggg atcatcaggt
2281 ttatccaatt ctcacacacc tgtgagacct ccaagtactt ctagtactgg cagtcgaggg
2341 agctgtgggt catcaggaag aactgctgag aagacaagtc ttagtttcaa atctgatcag
2401 gtgaagggtc agcaagaacc tgggactgaa gatgaaatat gtagcttttc aggaggtgta
2461 aaacaagaaa aaacagagga tggcaggagg agtgcttgca tgttgagcag tctgagagt
2521 agcttgacac cacctctctc aaccaacctg catctagaaa gtgaattgga tgcattggca
2581 agcctggaaa accatgtgaa aactgaacct gcagatatga atgaaagctg caaacagtca
2641 gggctcagca gccttgtaa tggaaagtcc ccaattcgaa gcctcatgca caggtcggca
2701 aggattggag gagatggcaa caataaagat gatgaccaa atgaagactg gtgtgctgtc
2761 tgccaaaacg gaggagatct cttgtgctgc gaaaaatgtc caaaggctt tcatctaact
2821 tgtcatgttc caacactact tagctttcca agtggggact ggatatgcac atttttaga
2881 gatattggaa agccagaagt tgaatatgat tgtgataatt tgcaacatag taagaagggg
2941 aaaactgcgc aggggttaag ccccgaggac caaaggaaat gtgaacgtct tctgctttac
3001 ctctattgcc atgaattaag tattgaattc caggagcctg ttctgtctc gataccaaac
3061 tactataaaa ttataaagaa accaatggat ttatccaccg tgaaaaagaa gcttcagaaa
3121 aaacattccc aacactacca aatcccgat gactttgtgg ccgatgtccg tttgatcttc
3181 aagaactgtg aaacctttaa tgaaatgatg aaagtgtgtc aagtttatgc agacacacaa

```

```
3241 gagattaatt tgaaggctga ttcagaagta gctcaggcag ggaaagcagt tgcattgtac
3301 tttgaagata aactcacaga gatctactca gacaggacct tcgcaccttt gccagagttt
3361 gagcaggaag aggatgatgg tgaggtaact gaggactctg atgaagactt tatacagccc
3421 cggcagaaaa cgcctaaagt cagatgagag accagtacat ataaagtaaa atgacatgga
3481 tttaaatgaa ttgttttaaaa aaaaaaaaaa
```

FIGURE 69

Figure 70

Name: TRAF2Gen Bank Accession No. U12597SEQ ID No.71

```

1  gaattccggc gcgctgcgac cgttggggct ttgttcgcgg gggtcacagc tctcatggct
61  gcagctagcg tgaccccccc tggctccctg gagttgctac agcccggtt ctccaagacc
121 ctccctggga ccaagctgga agccaagtac ctgtgctccg cctgcagaaa cgtcctccgc
181 aggcccttcc aggcgcagtg tggccaccgg tactgtcctt tctgcctggc cagcatcctc
241 agctctgggc ctcagaactg tgcctgctgt gttcacgagg gcatatatga agaaggcatt
301 tctatttttag aaagcagttc ggccttccca gataatgctg cccgcaggga ggtggagagc
361 ctgccggcgc tctgtcccag tgatggatgc acctggaagg ggaccctgaa agaatacgag
421 agctgccacg aaggccgctg cccgctcatg ctgaccgaat gtcccgctg taaaggcctg
481 gtccgccttg gtgaaaagga gcgccacctg gagcacgagt gcccggagag aagcctgagc
541 tgccggcatt gccgggcacc ctgctgcgga gcagacgtga aggcgcacca cgaggtctgc
601 cccaagttcc ccttaacttg tgacggctgc ggcaagaaga agatcccccg ggagaagttt
661 caggaccacg tcaagacttg tggcaagtgt cgagtccctt gcagattcca cgccatcggc
721 tgcctcgaga cggtagaggg tgagaaacag caggagcacg aggtgcagtg gctgcgggag
781 cacctggcca tgctactgag ctcggtgctg gaggcaaagc cctcttggg agaccagagc
841 cacgcggggt cagagctcct gcagaggtag gagagcctgg agaagaagac ggccactttt
901 gagaacattg tctgcgtcct gaaccgggag gtggagaggg tggccatgac tgccgaggcc
961 tgacgcccgc agcaccggct ggaccaagac aagattgaag cctgagtag caaggtgcag
1021 cagctggaga ggagcatttg cctcaaggac ctggcgatgg ctgacttggg gcagaaggtc
1081 aggcccttcc aggcgcagtg tggccaccgg tactgtcctt tctgcctggc cagcatcctc
1141 aggaagctcc aggaagctgt ggctggccgc ataccgcca tcttctcccc agccttctac
1201 accagcaggt acggctacaa gatgtgtctg cgtatctacc tgaacggcga cggcaccggg
1261 cgaggaacac acctgtccct cttctttgtg gtgatgaagg gcccgaatga cgccctgctg
1321 cggtagccct tcaaccagaa ggtgacctta atgctgctcg accagaataa cggggagcac
1381 gtgattgacg ccttcaggcc cgacgtgact tcatcctctt ttcagaggcc agtcaacgac
1441 atgaacatcg caagcggttg cccctctctt tgccccgtct ccaagatgga ggcaaagaat
1501 tcctacgtgc gggacgatgc catcttcata aaggccattg tggacctgac agggctctaa
1561 ctgcccccta ctggtgtctg ggggttgggg gcagccaggc acagccggct caccgagggg
1621 ccaccacgct gggccagggt ctactgtac aagtgggcag gggccccgct tgggcgcttg
1681 ggaggtgtgc ggcctgcagc caagttcact gtcacggggg aaggagccac cagccagtcc
1741 tcagatttca gagactgcgg aggggcttgg cagacggtct tagccaaggg ctgtggtggc
1801 attggccgag ggtcttcggg tgcctcccag cacaagctgc ccttgtgtgc ctgtgcagtg
1861 aaggagagag ccctgggttg gggacactca gagtgggagc acatcccagc agtgcccatg
1921 tagcaggagc acagtggatg gccttgtgtc cctcgggcat gacaggcaga aacgagggct
1981 gctccaggag aagggcctcc tgcctggccag agcaaggag gctgagcagc ttggttctcc
2041 cctctggccc ctggagagaa gggagcattc ctagaccctt ggtgcttgt ctgcacagag
2101 ctctggtctg tgccaccttg gccaggctgg ctgtgggagg gtctggtccc acgccgctc
2161 tgctcagaca ctgtgtggga gggcacagca cagctgcggg taaagtgtga gagcttgcca
2221 tccagctcac gaagacagag ttattaaacc attacaaatc tc

```

FIGURE 70

Figure 71

Name: TRAF3

Gen Bank Accession No. U21092

SEQ ID No.

72

```
1  cgggggagcg cggcgcgggc gccgcgtgcg cgagccgggg ttgcagccca gccgggactt
61  tccagccggc ggcagccgcg gcggtcgtcg gctcttcccc gcccccgtc atggggcagc
121 cgggggagca gaacgctgcg gaccgcggcg gaggacgcgc ccggcgcccc tgagccggcc
181 gagcggcgac ggaccgcgag aactcctctt tctaaaaatg gagtcgagta aaaagatgga
241 ctctcctggc gcgctgcaga ctaacccgcc gctaaagctg cacactgacc gtagtgctgg
301 gacgccagtt tttgtccctg aacaaggagg ttacaaggaa aagtttgatg agaccgtgga
361 ggacaagtac aagtgtgaga agtgccacct ggtgctgtgc agcccgaagc agaccgagtg
421 tgggcaccgc ttctgcgaga gctgcatggc ggccctgctg agctcttcaa gtccaaaatg
481 tacagcgtgt caagagagca tcgttaaaga taagggtgtt aaggataatt gctgcaagag
541 agaaattctg gctcttcaga tctattgtcg gaatgaaagc agaggttgtg cagagcagtt
601 aacgctggga catctgctgg tgcattttaa aaatgattgc ctttttgaag aacttccatg
661 tgtgcgtcct gactgcaaag aaaaggctct gaggaaagac ctgcgagacc acgtggagaa
721 ggcgtgtaaa taccgggaag ccacatgcag ccactgcaag agtcagggtc cgatgatcgc
781 gctgcagaaa cacgaagaca ccgactgtcc ctgcgtggtg gtgtcctgcc ctcaaatg
841 cagcgtccag actctcctga ggagcgagtt gagtgcacac ttgtcagagt gtgtcaatgc
901 ccccgagcacc tgtagtttta agcgctatgg ctgcgttttt caggggacaa accagcagat
961 caaggcccac gaggccagct ccgcccgtga gcacgtcaac ctgctgaagg agtggagcaa
1021 ctgcgtcgaa aagaaggttt ccttggtgca gaatgaaagt gtagaaaaaa acaagagcat
1081 acaaagtttg cacaatcaga tatgtagctt tgaaattgaa attgagagac aaaaggaaat
1141 gcttcgaaat aatgaatcca aaatccttca ttacagcga gtgatcgaca gccaagcaga
1201 gaaactgaag gagcttgaca aggagatccg gcccttccgg cagaactggg aggaagcaga
1261 cagcatgaag agcagcgtgg agtcctcca gaaccgcgtg accgagctgg agagcgtgga
1321 caagagtgcg gggcaagtgg ctcggaacac aggcctgctg gagtcccagc tgagccggca
1381 tgaccagatg ctgagtgtgc acgacatccg cctagccgac atggacctgc gcttccaggt
1441 cctggagacc gccagctaca atggagtgc catctggaag attcgcgact acaagcggcg
1501 gaagcaggag gccgtcatgg ggaagaccct gtccctttac agccagcctt tctacactgg
1561 ttacttttgt tataagatgt gtgccagggt ctacctgaac ggggacggga tggggaaggg
1621 gacgcacttg tcgctgtttt ttgtcatcat gcgtggagaa tatgatgccc tgcttcttg
1681 gccgtttaag cagaaagtga cactcatgct gatggatcag gggtcctctc gacgtcattt
1741 gggagatgca ttcaagcccg accccaacag cagcagcttc aagaagccca ctggagagat
1801 gaatatcgcc tctggctgcc cagtctttgt ggcccaaact gttctagaaa atgggacata
1861 tattaaagat gatacaattt ttattaaagt catagtggat acttcggatc tgcccgatcc
1921 ctgataagta gctggggagg tggatttagc agaaggcaac tcctctgggg gatttgaacc
1981 ggtctgtctt cactgagggt ctgcgctca gaaaaggacc ttgtgagacg gaggaagcgg
2041 cagaaggcgg acgcgtgccg gcgggaggag ccacgcgtga gcacacctga cacgttttat
2101 aatagactag ccacacttca ctctgaagaa ttatttatcc ttcaacaaga taaatattgc
2161 tgtcagagaa ggttttcatt ttcattttta aagatctagt taattaaggt ggaaaacata
2221 tatgctaaac aaaagaaaca tgatttttct tccttaaact tgaacaccaa aaaaacacac
2281 acacacacac acgtggggat agctggacat gtcagcatgt taagtaaaag gagaatttat
2341 gaaatagtaa tgcaattctg atatcttctt tctaaaattc aagagtgcaa ttttgtttca
2401 aatacagtat attgtctatt ttttaaggcct ccaaaaaaaa aaaaaattcc ggccg
```

FIGURE 71

Figure 72

Name: HIP116

Gen Bank Accession No. L34673

SEQ

ID No. 73

```
1 attccccggg tctgactgga ctgcggcgga cttacctttc agtcgtgcgc tcctgatccg
61 gcgctcggaa ttgtccccg gcttcagggc tgcggggcct ggaaggaggc gtatcgaggc
121 ggctcgaaaa cgatccaggg gagccgaggc gctcctcttg tcatcccact cagcgccatg
181 tcctggatgt tcaagagggg tccagtttgg aagtacttgc agactgtcca gtatggagtt
241 catggaaatt ttccacgcct ctcatatcca actttctttc cacgttttga attccaagat
301 gttatccctc cagatgactt totaactagt gatgaagaag tagattccgt tttatttgga
361 agtttgagag gtcattgtgt tggactacgc tattacacgg gagtagttaa taataatgaa
421 atggttgcac tacaacgaga tctaataaac cttatgata agaatgcaat taaagtaaac
481 aatgtgaatg gaaatcaagt tggccattta aagaaagagc ttgcagggtgc tttggcctat
541 atcatggaca acaaattggc acaaattgaa ggggtagttc cttttggtgc aaacaatgct
601 tttaccatgc ctctgcatat gacttttttg ggaaaagaag aaaatagaaa agcgggttca
661 gatcagttga agaaacatgg atttaaattg ggtcctgcac caaaaacttt aggattcaat
721 ttggaaagtg gttggggctc tgggaagagc ggaccaagct atagtatgcc agtgcattgt
781 gcagtacaga tgacaactga acagcttaaa acagaatttg acaaattgtt tgaagattta
841 aaagaagatg ataaaacca tgaaatggaa ccagctgagg ctattgaaa accactgctt
901 ccacatcaaa aacaagctct agcttggatg gtgtcacggg aaaatagcaa agaacttcca
961 ccattctggg aacagcgaaa tgacttatac tataacacaa taacaaattt ttctgagaag
1021 gaccgaccag aaaatgtcca tggaggaatt ttagctgatg atatgggttt gggtaaaact
1081 cttacagcca ttgcagtaat ccttaccaac ttccatgatg gcagacctct tcctattgaa
1141 agagttaaaa agaacttact gaagaaggaa tataatgtta acgatgactc tatgaaactt
1201 ggaggaaaca ataccagtga aaaggcagat ggactaagca aagacgcac tagatgtagt
1261 gaacaacca gtatttcaga tatcaaggag aagagtaagt ttccatgctc agaattgtct
1321 acgtcccgc ccaaaagaag aaaaactgct gtccagtaca tagaaagcag tgattcagag
1381 gaaattgaaa caagtgaatt gccgcagaaa atgaaaggca aactgaaaaa tgtacagtct
1441 gaaactaaag gcagggcgaa agcaggatct tctaaggtta tagaagatgt ggcatttgca
1501 tgtgcattaa cttcatctgt tcctacaaca aaaaagaaaa tgttgaaaaa gggagcttgt
1561 gcagtggagg ggtcaaagaa aactgatgtt gaggagagac caagaacaac actgatcatc
1621 tgtccgcttt ctgtgttaag caactggatt gaccagtttg gacaacatc aaactcagat
1681 gtacacttga atttttatgt ttattatggt cctgatcgta ttagagaacc ggccttactt
1741 tcaaacacgg atattgtttt gactacgtat aatattttta ctcatgacta tggaaactaa
1801 ggagatagtc cattacatag cataaggttg ctaagagtga tcctggatga aggacatgcc
1861 atacgaaatc caaatgctca gcagacaaaa gctgtacttg acttagaatc agaaagaaga
1921 tgggttttga caggactacc aatccagaat tctttaaagg acttgtggtc tcttctttcc
1981 tttttaaaaa ttaaacattt tattgataga gaatggtggc atagaacaat acagcgtcct
2041 gtcacaatgg gagatgaagg aggacttagg cgtttacagt ccctaattaa aaatattaca
2101 cttagaagaa caaagacaag caaaattaaa ggaaaacctg ttttgagatt accagaacgt
2161 aaagtattta ttcagcacat tacactttca gatgaagaga gaaagattta tcagtctgtg
2221 aaaaatgaag gcagagccac tattggaagg tatttttaat aagggaactgt cctggcacat
2281 tatgcagatg tcctgggtct tttgcttaga ctgcggcaaa tttgttgcca tacttacctt
2341 cttacaaatg cagtgtcttc caatggcccc tcaggaaatg atacacctga agaactgaga
2401 aagaagttaa taaggaagat gaagttaatt ctgagctcag gttcagatga ggaatgtgca
2461 atttgccctg attctttaac agttcctgtg ataacacatt gtgcacatgt attttgtaaa
2521 cctgtatttt gccaaagcat tcagaatgag cagccacatg ctaaatgcc tttatgcaga
2581 aatgatatac atgaagataa tttattagaa tgtcctccag aagaattagc acgtgacagt
2641 gagaaaaagt ctgatatgga atggacatcc agttcaaaga ttaatgcgct aatgcacgca
2701 ttgactgact taagaaagaa gaatcccaac ataaaaagtt tggttgtttc tcagtttaca
2761 acattcctgt ctttaataga aataccactt aaagcctctg gatttgtgtt tactcgtttg
2821 gatggttcca tggcccaaaa gaaaagagtt gaatcaattc agtgttttca aaacttgaa
2881 gcaggatctc caactataat gcttctgtcc ttaaaagcag gtggagttgg tttgaatctg
2941 tctgcagctt ctcgagtgtt ttaaatggat ccagcctgga atcctgctgc tgaagatcag
3001 tgctttgaca gatgccatag acttggctcag aagcaagaag ttatcatcac aaaattcatt
3061 gtaaaggact ctgttgaaga aaatatgctg aaaatacaaa aaaaaagag agaacttgca
3121 gcaggagcct ttggaactaa aaaccaaatt gctgacgaaa tgaaacaagc caaaattaat
3181 gaaatcagaa cattaattga cttataattt gtgggatttt agtaagaaga ctactatatg
```

3241 tgagaggcgt gatatctgga tggaagttgg gctggatgat ctccaaagtc gtttcaactc
3301 ttaaagacat cttaatcctg aatgtaaaca attggtatgt gtttagaatc agaatttgat
3361 tttgaacttg agtaattcat ccttacagct atctgtagaa ttagtcatct tttttctt

FIGURE 72

Figure 73

Name: STAF50Gen Bank Accession No. X82200SEQ ID No.74

```

1 gaattcggca cgagctcttc tcccctgatt caagactcct ctgctttgga ctgaagcact
61 gcaggagttt gtgaccaaga acttcaagag tcaagacaga aggaagccaa gggagcagtg
121 caatggattt ctcagtaaag gtagacatag agaaggaggt gacctgcccc atctgcctgg
181 agctcctgac agaacctctg agcctagatt gtggccacag cttctgccaa gcctgcatca
241 ctgcaaagat caaggagtca gtgatcatct caagagggga aagcagctgt cctgtgtgtc
301 agaccagatt ccagcctggg aacctccgac ctaatcggca tctggccaac atagttgaga
361 gagtcaaaga ggtcaagatg agcccacagg aggggcagaa gagagatgtc tgtgagcacc
421 atggaaaaaa actccagatc ttctgtaagg aggatggaaa agtcatttgc tgggtttgtg
481 aactgtctca ggaacaccaa ggtcaccaa cattccgcat aaacgaggtg gtcaaggaat
541 gtcaggaaaa gctgcaggta gccctgcaga ggctgataaa ggaggatcaa gaggctgaga
601 agctggaaga tgacatcaga caagagagaa ccgcttgaa gatcgagaga cagaagattc
661 tgaaaggggt caatgaaatg agagtcatct tggacaatga ggagcagaga gagctgcaaa
721 agctggagga aggtgaggtg aatgtgctgg acaacctggc agcagctaca gaccagctgg
781 tccagcagag gcaggatgcc agcacgtca tctcagatct ccagcggagg ttgacgggat
841 cgtcagtaga gatgctgcag gatgtgattg acgtcatgaa aaggagtga agctggacat
901 tgaagaagcc aaaatctgtt tccaagaaac taaagagtgt attccgagta ccagatctga
961 gtgggatgct gcaagttctt aaagagctga cagatgtcca gtactactgg gtggacgtga
1021 tgctgaatcc aggcagtgcc acttcgaatg ttgctatttc tgtggatcag agacaagtga
1081 aaactgtacg cacctgcaca tttaagaatt caaatccatg tgatttttct gcttttgggtg
1141 tcttcggctg ccaatatattc tcttcgggga aatattactg ggaagtagat gtgtctggaa
1201 agattgcctg gatcctgggc gtacacagta aaataagtag tctgaataaa aggaagagct
1261 ctgggtttgc ttttgatcca agtgtaaatt attcaaaagt ttactccaga tatagacctc
1321 aatatggcta ctgggttata ggattacaga atacatgtga atataatgct tttgaggact
1381 cctcctcttc tgatoccaa gttttgactc tctttatggc tgtgctccct gtcgtattgg
1441 ggttttccta gactatgagg caggcattgt ctcatttttc aatgtcacia accacggacg
1501 actcatctac aagttctctg gatgtcgctt ttctcgacct gcttatccgt atttcaatcc
1561 ttggaactgc ctagtcccca tgactgtgtg cccaccgagc tcctgagtgt tctcattcct
1621 ttaccacttt ctgcatagta gcccttctgt gagactcaga ttctgcacct gagttcatct
1681 ctactgagac catctcttcc tttctttccc cttcttttac ttagaatgtc ttgtattca
1741 ttgtctaggg cttccatagc aaagcatcat agattgctga tttaaactgt aattgtattg
1801 ccgtactgtg ggctgaaatc ccaaactag attccagcag agttggttct ttctgaggtc
1861 tgcaagggaag ggctctgttc catgcctctc tccttggtt gtagaaggca tcttgtccct
1921 atgactcttc acattgtctt tatgtacatc tctgtgcccc agttttccct ttttattaag
1981 acaccagtca tactggcctc agggcccacc gctaatgcct taatgaaatc attttaacat
2041 tatattgtgt acaaagacct tatttccaaa taagataata tttggaggta ttgggaataa
2101 aatttgagga aggcgatttc actcataaca atcttaccct ttcttgcaag agatgcttgt
2161 acattatttt cctaatacct tggtttctact agtagtaaac attattattt tttttatatt
2221 tgcaaaggaa acatatctaa tccttcctat agaaagaaca gtattgctgt aattcctttt
2281 cttttcttcc tcatttcttc tgccccttaa aagattgaag aaagagaaac ttgtcaactc
2341 atatccacgt tatctagcaa agtcataaga atctatcact aagtaatgta tccttcagaa
2401 tgtgttgggt taccagtgc accccatatt catcacaaaa ttaaagcaag aagtcctatg
2461 taatttattt gctaatagtg gatttttaac gctcagagtt tctgaggtca aattttatct
2521 ttctacttac aagctctatg atcttaataa atttacttaa tgtattttgg tgtattttcc
2581 tcaaattaat attggtgttc aagactatat ctaattcctc tgatcacttt gagaaacaaa
2641 cttttattaa atgtaaggca cttttctatg aattttaaat ataaaaataa atattgttct
2701 gattattact gaaaagatgt cagccatttc aatgtcttgg gaaacaattt tttgtttttg
2761 ttctgttttc tttttgcttc aataaaacaa tagctggctc taaaaaaaaa a

```

FIGURE 73

Figure 74

Name: HLTF-1
ID No. 75

Gen Bank Accession No. Z46606

SEQ

```

1  ttgcctctc tcatgatccc aggaagacta acaggctctt tgtgacactg gcactgcgct
61  gtcggatcca gctgttgcca ttgtgaaggg gtgttctctt cctcatttcc tgaaggattt
121 gttgcagaag gagacggcgt cgacgtctga ctggactcgc ggcgacttac ctttcagtcg
181 tgcgctctcg atccggcgct cgggaatttg ccccggttc agggctgcgg ggccgtggaag
241 gaggcgtatc gaggcggctc gaaaacgac caggggagcc gaggcgctcc tcttgtcatc
301 ccactcagcg ccatgtcctg gatgttcaag agggatccag tttggaagta cttgcagact
361 gtccagtatg gagttcatgg aaattttcca cgcctctcat atccaacttt ctttctacgt
421 tttgaattcc aagatgttat cctccagat gactttctaa ctagtatga agaagtagat
481 tccgttttat ttggaagttt gagaggcat gtggttggac tacgctatta cacgggagta
541 gttataata atgaaatggg tgcattacaa cgagatccta ataaccctta tgataagaat
601 gcaattaaag taaacaatgt gaatggaaat caagtgggcc atttaaagaa agagcttgca
661 ggtgctttgg cctatatcat ggacaacaaa ttggcacaaa ttgaaggggt agttcctttt
721 ggtgcaaaac atgcttttac catgcctctg catatgactt tttggggaaa agaagaaaat
781 agaaaagcgg tttcagatca gttgaagaaa catggattta aattgggtcc tgcacaaaaa
841 actttaggat tcaatttggg aagtgggtgg ggctctggaa gagctggacc aagctatagt
901 atgccagtgc atgctgcagt acagatgaca actgaacagc ttaaaacaga atttgacaaa
961 ttgtttgaag atttaaaaga agatgataaa acccatgaaa tggaaaccagc tgaggctatt
1021 gaaacaccac tgcctccaca tcaaaaacaa gctctagctt ggatgggtgc acgggaaaat
1081 agcaaagaac tccaccatt ctgggaacag cgaaatgact tatactataa cacaataaca
1141 aatttttctg agaaggaccg accagaaaat gtccatggag gaattttagc tgatgatatg
1201 ggtttgggta aaactcttac ggccattgca gtaatcctta ccaacttcca tgatggcaga
1261 cctcttctta ttgaaagagt taaaaagaat ctactgaaga aggaatataa tgtaaacgat
1321 ggctctatga aacttggagg aaacaatacc agtgaagagg cagatggact aagcaaagac
1381 gcatctagat gtagtgaaca acccagtatt tcagatatca aggagaagag taagtttcgc
1441 atgtcagaat tgtctagctc ccgccccaaa agaagaaaaa ctgctgtcca gtacatagaa
1501 agcagtgatt cagaggaaat tgaaacaagt gaattgccgc agaaaatgaa agggaaactg
1561 aaaaatgtac agtctgaaac taaaggcagg gcgaaagcag gatcttctaa ggttatagaa
1621 gatgtggcat ttgcatgtgc attaacttca tctgttctta caacaaaaaa gaaaatgttg
1681 aaaaagggag cttgtgcagt ggaggggtca aagaaaactg atgttgagga gagaccaaga
1741 acaacactga tcatctgtcc gctttctgtg ttaagcaact ggattgacca gtttgacaa
1801 catataaaat cagatgtaca cttgaatttt tatgtttatt atggctctga tcgtattaga
1861 gaaccggcct tactttcaaa acaggatatt gttttgacta cgtataatat tttaactcat
1921 gactatggaa cttaaaggaga tagtccatta catagcataa ggtggctaag agtgatcctg
1981 gatgaaggac atgccatacg aaatccaaat gctcagcaga caaaagctgt acttgactta
2041 gaatcagaaa gaagatgggt tttgacaggt actccaatcc agaattcttt aaaggacttg
2101 tggctctctc tttccttttt aaaacttaaa ccatattattg atagagaatg gtggcataga
2161 acaatacagc gtccgtgtcac aatgggagat gaaggaggac ttaggcgttt acagtcccta
2221 attaaaaata ttacacttag aagaacaaag acaagcaaaa ttaaaggaaa acctgttttg
2281 gagttaccag aacgtaaagt atttattcag cacattacac tttcagatga agagagaaag
2341 atttatcagt ctgtgaaaaa tgaaggcaga gccactattg gaaggatttt taatgaaggg
2401 actgtcctgg cacattatgc agatgtcctg ggtcttttgc ttagactgcg gcaaatttgt
2461 tgccatactt accttcttac aaatgcagtg tcttccaatg gccctcagg aaatgataca
2521 cctgaagaac tgagaaagaa gttaataagg aagatgaagt taattctgag ctcaggttca
2581 gatgaggaat gtgcaatttg cctggattct ttaacagttc ctgtgataac acattgtgca
2641 catgtatttt gtaaaccctg tatttgccaa gtcattcaga atgagcagcc acatgctaaa
2701 tgccctttat gcagaaatga tatacatgaa gataatttat tagaatgtcc tccagaagaa
2761 ttagcacgtg acagtgagaa aaagtctgat atggaatgga catccagttc aaagattaat
2821 gcgctaattg acgcattgac tgacttaaga aagaagaatc ccaacataaa aagtttggtt
2881 gtttctcagt ttacaacatt cctgtcttta atagaaatac cacttaaagc ctctggattt
2941 gtgtttactc gtttggatgg ttccatggcc caaaagaaaa gagttgaatc aattcagtgt
3001 tttcaaaaca ctgaagcagg atctccaaat ataatgcttc tgccttaag agcagtgga
3061 gttggtttga atctgtctgc agcttctcga gtgtttttta tggatccagc ctggaatcct
3121 gctgctgaag atcagtgtt tgacagatgc catagacttg gtcagaagca agaagttatc
3181 atcacaaaat tcattgtaaa ggactctgtt gaagaaaaata tgctgaaaat acaaaacaaa

```

3241 aagagagaac ttgcagcagg agccttttga actaaaaaac caaatgctga cgaaatgaaa
3301 caagccaaaa ttaatgaaat cagaacatta attgacttat aatttgtggg attttagtaa
3361 ggtcagtttg attggatact taagtttttag aaatgagaaa aatacagagt tttagaaatg
3421 agatctagag aacacgtctt ctaaaagggg catattttat attagtgaag aggtattact
3481 gacacaattt cttctatata tgaacctatt tttaatgaaa cttcaaatag caataagttc
3541 cgttatatac tgtggcctga aataatttga gaaaaaagg tactttgtta ttcagctttt
3601 cataatatct atgctgagta ttttcacgta tcttccaagt actcagcttt ttcgtatttc
3661 aaataagggtc agacctttta tacttttgac caaatagtta atttctatgt tggacactta
3721 gttattttacc aaagcctccg atttgtgatg cagtgtttgt agtccttgta aacaatatat
3781 acagactata caagaattaa ttttattggg ctttcaaaaa ccatatttgc attccagaac
3841 caaatcttaa atgagaccaa agtccaggtt agcacaggtt tttatttttc ctacagctac
3901 attgagatat agttcacata aatgacttgg gagttttatg ttcataaaaa aattagggat
3961 tatgttaaga gtactatttt ttccattttta gtaaagtagt actacacaca cattgtttta
4021 atgtaactgc tcgtgtctga ggtataaata tagtctgtgg gagtgagagc aaaccagtcc
4081 taccacaattt gatttgaata ttttaaatta tggaacctgc taaatagata tttctataaa
4141 tagataattt ttatttatgt agcttttttt ggaagtaact ttataaattt ttataattca
4201 gaagactact atatgtgaga ggcgtgatat ctggatggaa gttgggctgg atgatctcca
4261 aagtcgtttc aactcttaaa gacatcttaa tctgaatgt aaacaattgt tatgtgttta
4321 gaatcagaat ttgattttga acttgagtaa ttcatactta cagctatctg tagaattagt
4381 catctttttt cttttctttt ttttttactt ttttgttaat aagcaaaact atttgctgca
4441 gaatttgggt tgactcctga gcatatttaa aacaaagaag ctagaaaattt agcagtcaga
4501 ttaggtaggt ggttttattt caaagggaaa ctttaatcca aagaaagatt aattactcta
4561 acaaacaaga gaagcttcat gtttgatgat acagatttaa gaatacctga gacttaagag
4621 tgttggaagt catttgatag aaaagatgag atcggagaca atgttgtgtt atagggcaca
4681 cattgaaggt atatgccaaa tctctcacia gataagtcct tttctccctg tgcggttttc
4741 ttgctcaaag ggaacagtga attagccagc tagaatcttc ctggtccctt tttgaggcag
4801 tagcaggtta ggaaagggct gattttcatc aaaaccaaga cctttctgca ggcattgatag
4861 tggaataata atgtgggatt agcccgttag cattaggagc agttgggaag ttacctggta
4921 gatcaagcat tacacacaaa aaatcaagtt gatcagagta tgggttctgc atgtagcaat
4981 acttcagtga gattaagtat aaacagtttt tggcaaaaaa caacacagtc tactctttct
5041 gcttacaaaag acaaagcctt acaaactcac tatgaagggt aaagggagac agotttgctt
5101 tttgcccaga catttacaaa gttgttttta aaacacactc ataagtaagt ttggcaagtt
5161 gtttaaaaaa tgtctctttg ttttgtacag ttctgttaga tgttgttata ttttaaaagt
5221 ttaattttaa aaattttaatt tgtccttcct aagaaggata aatatataaa aaagccatct
5281 ggaatgaaaa cttcctatat gctagctgtt gtcttattat tatatagaaa aataacttta
5341 gaaaaatatt gaagacattg tattaccact tgtgattcaa acaattttgt ggttaaaact
5401 ggatttttaa tttaaaaatc aataaaaatt tcaaatgtta aaaaaaaaaa aaaaaaa

FIGURE 74

Figure 75

Name: BIR4Gen Bank Accession No. U45880SEQ ID No.76

```

1 gaaaagggtgg acaagtccta ttttcaagag aagatgactt ttaacagttt tgaaggatct
61 aaaacttgtg tacctgcaga catcaataag gaagaagaat ttgtagaaga gtttaataga
121 ttaaaaaactt ttgctaattt tccaagtggg agtcctgttt cagcatcaac actggcacga
181 gcagggtttc tttatactgg tgaaggagat accgtgcggg gcttttagttg tcatgcagct
241 gtagatagat ggcaatatgg agactcagca gttggaagac acaggaaagt atcccaaagt
301 tgcagattta tcaacggcctt ttatcttgaa aatagtgcc cgcagtctac aaattctggt
361 atccagaatg gtcagtacaa agttgaaaac tatctgggaa gcagagatca ttttgcctta
421 gacaggccat ctgagacaca tgcagactat cttttgagaa ctgggcagggt ttagatata
481 tcagacacca tataccgag gaaccctgcc atgtattgtg aagaagctag attaaagtcc
541 tttcagaact ggccagacta tgctcaccta accccaagag agttagcaag tgctggactc
601 tactacacag gtattgggtga ccaagtgcag tgcttttgtt gtgggtggaaa actgaaaaat
661 tgggaacctt gtgatcgtgc ctggtcagaa cacaggcgac actttcctaa ttgcttcttt
721 gttttgggcc ggaatcttaa tattcgaagt gaatctgatg ctgtgagttc tgataggaat
781 ttcccaaagt caacaaatct tccaagaaat ccatccatgg cagattatga agcacggatc
841 tttacttttg ggacatggat atactcagtt aacaaggagc agcttgcaag agctggattt
901 tatgctttag gtgaagggtga taaagtaaag tgctttcact gtggaggagg gctaactgat
961 tggaagccca gtgaagaccc ttgggaacaa catgctaaat ggtatccagg gtgcaaatat
1021 ctgttagaac agaagggaca agaatatata aacaatatc atttaactca ttcacttgag
1081 gagggtctgg taagaactac tgagaaaaca ccatcactaa ctagaagaat tgatgatacc
1141 atcttccaaa atcctatggg acaagaagct atacgaatgg gggtcagttt caaggacatt
1201 aagaaaataa tggaggaaaa aattcagata tctgggagca actataaatc acttgagggtt
1261 ctgggtgcag atctagtga tgcctcagaa gacagtatgc aagatgagtc aagtcagact
1321 tcattacaga aagagattag tactgaagag cagctaaggc gcctgcaaga ggagaagctt
1381 tgcaaaatct gtatggatag aaatattgct atcgtttttg ttccttgtgg acatctagtc
1441 acttgtaaac aatgtgctga agcagttgac aagtgtccca tgtgtacac agtcattact
1501 ttcaagcaaa aaatttttat gtcttaactc aactctatag taggcatggt atgttgttct
1561 tattaccctg attgaatgtg tgatgtgaac tgactttaag taatcaggat tgaattccat
1621 tagcatttgc taccaagtag gaaaaaaaat gtacatggca gtgttttagt tggcaatata
1681 atctttgaat ttcttgattt ttcagggtat tagctgtatt atccattttt tttactgtta
1741 ttttaattgaa accatagact aagaataaga agcatcatac tataactgaa cacaatgtgt
1801 attcatagta tactgattta atttctaagt gtaagtgaat taatcatctg gattttttat
1861 tcttttcaga taggcttaac aaatggagct ttctgtatat aaatgtggag attagagtta
1921 atctccccaa tcacataatt tgttttgtgt gaaaaaggaa taaattgttc catgctgggtg
1981 gaaagataga gattgttttt agaggttggg tgttgtgttt taggattctg tccattttct
2041 tgtaaaggga taaacacgga cgtgtgcgaa atatgtttgt aaagtgattt gccattgttg
2101 aaagcgtatt taatgataga atactatcga gccaacatgt actgacatgg aaagatgtca
2161 gagatatgtt aagtgtaaaa tgcaagtggc gggacactat gtatagtctg agccagatca
2221 aagtatgtat gttgttaata tgcatagaac gagagatttg gaaagatata caccaaactg
2281 ttaaatgtgg tttctcttcg gggagggggg gattggggga ggggccccag aggggtttta
2341 gaggggcctt ttcactttcg acttttttca ttttgttctg ttcggatttt ttataagtat
2401 gtagaccccg aagggtttta tgggaactaa catcagtaac ctaacccccg tgactatcct
2461 gtgctcttcc tagggagctg tgttgtttcc caccaccac ccttcctctt gaacaaatgc
2521 ctgagtgtct gggcactttg

```

FIGURE 75

Figure 76

Name: ADENOVIRUS 5Gen Bank Accession No. X86098SEQ ID No.77

E1A-Binding Protein

```

1 ggagcataat gctaaagaag taaacaggtc atggcacggt taacaaaaag acgacaggcg
61 atacaaaaagc tatccagcat ctttgggcag ccattgagat tatacggaac cagaagcaga
121 ttgccaaacat tgaccgtatt acaaaatgtg aaacaactac attattcttg aacctatggg
181 gattttttaca tcattacaca gatatgtcat ttccattagt tgtatcattg ttataaactg
241 gtatatgtct cgagtccacg gtatgcaccc taaagagacc acccgctcagc tgagccttagc
301 tgtgaaagat ggtcttattg tcgaaactct aacagtgggc tgcaaagggt caaaagctgg
361 tattgaacaa gaaggatatt ggttgccagg agatgagatt gactgggaaa cagaaaatca
421 tgactgggtat tgttttgaat gccatttgcc tggagagggt ttgatattgt acctgtgttt
481 tcgtgtgtat cattccaagt gtttgtctga tgagttcagg cttagagaca gcagtagtcc
541 ctggcagtgcc ccagtttgca ggagcattaa gaagaagaat acaaacaaac aggagatggg
601 cacatacctc agattcattg tctcccgcat gaaggagagg gctatagatc ttaataaaaa
661 ggggaaggac aataaacacc cgatgtacag gaggctgggt cactcagctg tggacgttcc
721 caccattcaa gagaaagtga atgaaggga ataccgaagt tatgaagagt tcaaagctga
781 tgcccaattg cttctccaca ataccgtgat tttctatgga gcagacagtg agcaagctga
841 cattgcgagg atgctatata aagacacatg tcatgagctg gatgaactgc agctttgcaa
901 gaattgcttt tacttgtcaa atgctcgtcc tgacaactgg ttctgttatc cttgtatacc
961 taatcatgag ctggtttggg ctaaaatgaa aggttttggg ttttggccag ccaaagtcac
1021 gcagaaagaa gacaatcaag tcgacgttcg cttctttggc caccaccacc agagggcctg
1081 gattccttct gaaaacattc aagatatcac agtcaacatt catcggtctg acgtgaagcg
1141 cagtatgggt tggaaaaagg cctgtgatga gctggagctg catcagcggt tcctacgaga
1201 agggagattt tggaaatcta agaattgagga ccgaggtgag gaagaggcag aatccagtat
1261 ctctccacc agtaatgagc agctaaaggt cactcaagaa ccaagagcaa agaaaggacg
1321 acgtaatcaa agtgtggagc ccaaaaagga agaaccagag cctgaaacag aagcagtaag
1381 ttctagccag gaaataacca cgatgcctca gcccatcgaa aaagtctccg tgtcaactca
1441 gacaaagaag ttaagtgcct cttcaccaag aatgctgcat cggagcacc agaccacaaa
1501 cgacggcgtg tgtcagagca tgtgccatga caaatacacc aagatcttca atgacttcaa
1561 agaccggatg aagtccgacc acaagcggga gacagagcgt gttgtccgag aagctctgga
1621 gaagctgcgt tctgaaatgg aagaagaaaa gagacaagct gtaataaaag ctgtagccaa
1681 catgcagggt gagatggaca gaaaatgtaa gcaagtaaag gaaaagtgt aggaggaatt
1741 tgtagaagaa atcaagaagc tggcaacaca gcacaagcaa ctgatttctc agaccaagaa
1801 gaagcagtggt tgctacaact gtgaggagga ggccatgtac cactgctgct ggaacacatc
1861 ctactgctcc atcaagtgcc agcaggagca ctggcacgag gagcacaagc gcacctgccg
1921 ccggaaaaaga tgaagctggc cttcccgga gtcacccga tgattactct tttcagacac
1981 agcgggtttt gtttccaaga agccaaaatt gtttagaatt tgcttcccat tttgcaccag
2041 cctttaaaca cttttcgtga agaaattttg cacagtagtt taaatctttt gttaatgctc
2101 ctccgaagtt tttcaggggg taaaagtaac atcagtggag ggtattatth taaataaatt
2161 ttaattgaga atttgttgca ttttcagcaa attttaaaac atttttaggt tttacagaga
2221 ttttaacctt taaacaacag atctttaaaa aacaggtgaa tacaagtgag ttaacaaaag
2281 aaacatttag aatagatctg aatgtaagaa ctacagaact gtttcagaaa taaaacatac
2341 taccttgatg tgacattttt ttcttaacct tgttgagctg gttttgttca gcttaattta
2401 ctgttcaaag gcattatctg ttggtcacac cagtgggtat atgattgaat ttagggaaca
2461 gggttgacac agcagggcta gtctgcata ttttttctta aatatttccc aattgtgttt
2521 ttcattatth cttttcaata tataactttt ataacaaatt attagctttg atctttagt
2581 ttaaaattgc agggaactgg ggtaatcttt tactgagctg gatcttagag aaaatgaata
2641 tttaaatttt aaagtgtgcc acatttcac tttgtcctaa catgagtgtc tgaacaaaa
2701 taaacaaca aaaacaaagc ct

```

FIGURE 76

Figure 77

Name: HT2AGen Bank Accession No. U18543SEQ ID No.78

```
1  gtgggctcgt cggagccgcg ggcggtcagc aggaatttga ccctctaggg catgaatact
61  gtgctgttca gttctgagct gtgctagcaa tacccttcaa aggaagagca atggctgcag
121  cagcagcttc tcacctgaac ctggatgccc tccgggaagt gctagaatgc cccatctgca
181  tggagtccat cacagaagag cagctgcgtc ccaagcttct gcactgtggc cataccatct
241  gccgccagtg cctggagaag ctattggcca gtagcatcaa tgggtgtccg tgtccctttt
301  gcagcaagat taccgcata accagcttga cccagctgac agacaatctg acagtgctaa
361  agatcattga tacagctggg ctacgcgagg ctgtggggct gctcatgtgt cggctctgtg
421  ggcggcgctc gccccggcaa ttctgccgga gctgtgggtt ggtgttatgt gagccctgcc
481  gggaggcaga ccatcagcct cctggccact gtacactccc tgtcaaagaa gcagctgagg
541  agcggcgctc ggacttttga gagaagttaa ctctgtctgc ggaacttatg ggggagctgc
601  agcggcggaa ggcagccttg gaaggtgtct ccaaggacct tcaggcaagg tataaagcag
661  ttctccagga gtatgggcat gaggagcgca ggtccagga tgagctggct cgctctcgga
721  agttcttcac aggtcttttg gctgaagttg agaagtccaa tagtcaagtg gtagaggagc
781  agagttacct gcttaacatt gcagaggtgc aggtgtgtc tcgctgtgac tacttcctgg
841  ccaagatcaa gcaggcagat gtagcactac tggaggagac agctgatgag gaggagccag
901  agctcactgc cagcttgctc cgggagctca cctgcaaga tgtggagctc ctttaaggtag
961  gtcattgttg cccctccaa attggacaag ctgttaagaa gccccggaca gtttaacgtg
1021 aagattcctg ggccatggag gccacagcgt ctgctgcctc tacctctgtt acttttagag
1081 agatggacat gagcccgag gaagtggttg ccagccctag ggctcacct gctaaacagc
1141 ggggtcctga ggcagcctcc aatatccagc agtgcctctt tctcaagaag atgggggcca
1201 aaggcagcac tccaggaatg ttcaatcttc cagtcagtct ctacgtgacc agtcaagggtg
1261 aagtactagt cgctgaccgt ggtaactatc gtatacaagt ctttaccgcg aaaggctttt
1321 tgaaggaaat ccgccgcagc cccagtggca ttgatagctt tgtgctaagc ttccttgggg
1381 cagatctacc caacctcact cctctctcag tggcaatgaa ctgccagggg ctgattgggtg
1441 tgactgacag ctatgataac tccctcaagg tatatacctt ggatggccac tgcgtggcct
1501 gtcacaggag ccagctgagc aaaccatggg gtatcacagc cttgccatct ggccagtttg
1561 tagtaaccga tgtggaaggt ggaaagcttt ggtgtttcac agttgatcga ggatcagggg
1621 tgggtcaaata cagctgccta tgtagtgtct tgccggccaa atttgtcacc tgtgatgctg
1681 agggcaccgt ctacttcacc cagggttag gcctcaatct ggagaatcgg cagaatgagc
1741 accacctgga ggggtggctt tccattggct ctgtaggccc tgatgggcag ctgggtcgcc
1801 agattagcca cttcttctcg gagaatgagg atttccgctg cattgctggc atgtgtgtgg
1861 atgctcgtgg tgatctcatc gtggctgaca gtagtcgcaa ggaaattctc cattttctta
1921 agggcggggg ctatagtgtc cttattcgag agggacttac ctgtccggtg ggcataagccc
1981 taactcctaa ggggcagctg ctggtcttgg actgttggga tcattgcacg aagatctaca
2041 gctaccatct gagaagatat tccaccccat aggggatgag aaattatcag tttcttctgc
2101 tccaagcca acttcccttc ccttagttct tgggtgttag tggcacatgc agaataagact
2161 cagcctatgt cctgattcca gctgggtagt tctagaactt tagaagctcc atcttttaat
2221 gtttttattt gttatgtccc cctcccggct tcccacctaa atttagagct ttaaaagatg
2281 cactgcccga ataggacaca cgatgggtgt agctgaagtt tgattagcaa ttaggcactt
2341 ccaaggcttt agtagagaga gccactttag ccctttgtgc catgtttgaa atttgccctt
2401 gtattaaatc cttgattttt tccc
```

FIGURE 77

Figure 78

Name: TRAF3

Gen Bank Accession No. U21092

SEQ ID No.

79

```
1  cgggggagcg cggcgcgccc gccgcgtgcg cgagccgggg ttgcagccca gccgggactt
61  tccagccggc ggcagccgcg gcggtcgteg gctcttcccc gcccccgtc atggggcagc
121 cgggggagca gaacgctgcg gaccgcggcg gaggacgcgc ccggcgcccc tgagccggcc
181 gagcggcgac ggaccgcgag aactcctctt tcctaaaatg gagtcgagta aaaagatgga
241 ctctcctggc gcgctgcaga ctaaccgcgc gctaaagctg cacactgacc gtagtgctgg
301 gacgccagtt tttgtccctg aacaaggagg ttacaaggaa aagtttgtga agaccgtgga
361 ggacaagtac aagtgtgaga agtgccacct ggtgctgtgc agcccgaaac agaccgagtg
421 tgggcaccgc ttctgcgaga gctgcatggc ggccctgctg agctcttcaa gtccaaaatg
481 tacagcgtgt caagagagca tcgttaaaga taagggtgtt aaggataatt gctgcaagag
541 agaaattctg gctcttcaga tctattgtcg gaatgaaagc agaggttgtg cagagcagtt
601 aacgctggga catctgctgg tgcatttaaa aaatgattgc cattttgaag aacttccatg
661 tgtgcgtcct gactgcaaag aaaaggctct gaggaaagac ctgcgagacc acgtggagaa
721 ggcgtgtaaa taccgggaag ccacatgcag ccactgcaag agtcaggttc cgatgatcgc
781 gctgcagaaa cacgaagaca ccgactgtcc ctgcgtggtg gtgtcctgcc ctcaaatg
841 cagcgtccag actctcctga ggagcgagtt gagtgcacac ttgtcagagt gtgtcaatgc
901 cccagcacc tgtagtttta agcgtatggt ctgcgttttt caggggacaa accagcagat
961 caaggccac gaggccagct ccgcgtgca gcacgtcaac ctgctgaagg agtggagcaa
1021 ctgcctcgaa aagaaggttt ccttggtgca gaatgaaagt gtagaaaaaa acaagagcat
1081 acaaagtttg cacaatcaga tatgtagctt tgaaattgaa attgagagac aaaaggaaat
1141 gcttcgaaat aatgaatcca aaatccttca tttacagcga gtgatcgaca gccaagcaga
1201 gaaactgaag gagcttgaca aggagatccg gcccttccgg cagaactggg aggaagcaga
1261 cagcatgaag agcagcgtgg agtccctcca gaaccgcgtg accgagctgg agagcgtgga
1321 caagagtgcg gggcaagtgg ctcggaacac aggcctgctg gagtcccagc tgagccggca
1381 tgaccagatg ctgagtgtgc acgacatccg cctagccgac atggacctgc gcttccaggt
1441 cctggagacc gccagctaca atggagtgt catctggaag attcgcgact acaagcggcg
1501 gaagcaggag gccgtcatgg ggaagaccct gtccctttac agccagcctt tctacactgg
1561 ttactttggt tataagatgt gtgccagggt ctacctgaac ggggacggga tggggaaggg
1621 gacgcacttg tcgctgtttt ttgtcatcat gcgtggagaa tatgatgcc tgcttccttg
1681 gccgtttaag cagaaagtga cactcatgct gatggatcag gggtcctctc gacgtcattt
1741 gggagatgca ttcaagccg accccaacag cagcagcttc aagaagccca ctggagagat
1801 gaatatcgcc tctggctgcc cagtctttgt ggcccaaact gttctagaaa atgggacata
1861 tattaaagat gatacaattt ttattaaagt catagtggat acttcggatc tgcccgatcc
1921 ctgataagta gctggggagg tggatttagc agaaggcaac tcctctgggg gatttgaacc
1981 ggtctgtctt cactgaggtc ctgcgctca gaaaaggacc ttgtgagacg gaggaagcgg
2041 cagaaggcgg acgcgtgccg gcgggaggag ccacgcgtga gcacacctga cacgttttat
2101 aatagactag ccacacttca ctctgaagaa ttatttatcc ttcaacaaga taaatattgc
2161 tgtcagagaa ggtttttcatt ttcattttta aagatctagt taattaaggt ggaaaacata
2221 tatgctaaac aaaagaaaca tgatttttct tccttaaact tgaacaccaa aaaaacacac
2281 acacacacac acgtggggat agctggacat gtcagcatgt taagtaaaag gagaatttat
2341 gaaatagtaa tgcaattctg atatcttctt tctaaaattc aagagtgcaa ttttgtttca
2401 aatacagtat attgtctatt ttttaaggcct ccaaaaaaaaa aaaaaattcc ggccg
```

FIGURE 78

Figure 79: Human PRT3 full mRNA, Annotated Sequence

---- - gi|10432611|dbj|AK021429.1|AK021429 Homo sapiens cDNA FLJ11367
fis, clone HEMBA1000303, highly similar to Mus musculus Plenty of SH3s
(PRT3) mRNA

---- - gi|7959248|dbj|AB040927.1|AB040927 Homo sapiens mRNA for
KIAA1494 protein, partial cds

----- - Both hPOSH and KIAA1495

----- - Ring Domain

----- - SH3 Domian

----- - start codon and stop codon of predicted ORF

CTGAGAGACACTGCGAGCGGCGAGCGGGTGGGGCCGCATCTGCATCAGCCGCCGAGCCGCTGCGGGGG
CGCGAACAAAGAGGAGGAGCCGAGGCGGAGAGCAAAAGTCTGAAATGGATGTTACATGAGTCATTTTAAG
GGATGCACACAACATATGAACATTTCTGAAGATTTTTTCTCAGTAAAGTAGATAAAGATGGATGAATCAGC
CTTGTGGATCTTTTGGAG
GCAGGACTCTTGTGGCTCGGGTGTGAGGAGCTTCCCAGTAACATCTTGCTGGTCAGACTTCTGGATGG
CATCAAACAGAGGCCTTGGAAACCTGGTCCTGGTGGGGGAAGTGGGACCAACTGCACAAATGCATTAAGG
TCTCAGAGCAGCACTGTGGCTAATTGTAGCTCAAAAGATCTGCAGAGCTCCCAGGGCGGACAGCAGCCTC
GGGTGCAATCCTGGAGCCCCCAGTGAGGGGTATACCTCAGTTA
AACCGTTACCTCAGCCCCCA
AGCTGATAGAATGGGATAAGCCTCCTGTGCCAGGAGTTGATGCTGGAGAATGTTCTCGGCAGCAGCCCA
GAGCAGCACTGCCCCAAAGCACTCCGACACCAAGAAGAACACCAAAAAGCGGCACTCCTTCACTTCCCTC
ACTATGGCCAACAAGTCTCCAGGCATCCAGAACCGCCACTCCATGGAGATCAGCCCCCTGTCTCTCA
TCAGCTCCAGCAACCCCACTGTGCTGCACGGATCAGCGAGCTGTCTGGGCTCTCTGAGTGCCCCCTC
TCAGGTTTCATATAAGTACCACCGGGTTAATTGTGACCCCGCCCCAAGCAGCCAGTGACAACCTGGCCCC
TCGTTTACTTTCCCATCAGATGTTCCCTACCAAGCTGCCCTTGGAACTTTGAATCCTCTCTCCACCAC
CCCCCTCTCTGGCTGCCACTGTCTTGCCTCCACACCAECAGGCGCCACCGCCGCCGCTGTGCTGCTGG
AATGGGACCGAGGCCCATGGCAGGATCCACTGACCAGATTGCACATTTACGGCCGAGACTCGCCCC
ACAAGGGCGGTGACAAATGCTTCCCAAGCTAAAGTCCTATG
TCTACAGCTGGCCAGACAAGTCGGGGAGTGACCATGGTCAGTCCTTCCACGGCAGGAGGGCCTGCCCAGA
AGTCCAGGGAAATGGCGTGGCTGGGAGTCCAGTGTTGTCCCGCAGCTGTGGTATCAGCAGCTCACAT
CCAGACAAGTCTCAGGCTAAGGTCTTGTGTCACATGACGGGGCAAATGACAGTCAACCAGGCCCGCAAT
GCTGTGAGGACAGTTGACGCGCACAACCAGGAACGCCCCACGGCAGCAGTGACACCCATCCAGGTACAGA
ATGCCGCCGGCCTCAGCCCTGCATCTGTGGGCTGTCCCATCACTCGCTGGCCTCCCCACAACCTGCGCC
TCTGATGCCAGGCTCAGCCACGCACACTGCTGCCATCAGTATCAGTCGAGCCAGTGCCCTCTGGCCTGT
GCAGCAGCTGCTCCACTGACTTCCCCAAGCATCACCAGTGCTTCTCTGGAGGCTGAGCCCAGTGCCCGGA
TAGTGACCGTTCTCCCTGGACTCCCCACATCTCTGACAGTGCTTCATCAGCTTGTGGGAACAGTTTCAGC
AACCAAACAGACAAGGATAGCAAAAAGAAAAAAGGGTTTGTGAAGTTGCTTTCTGGCGCCTCCACT
AAACGGAAGCCCCGCGTGTCTCTCCAGCATCGCCACCCCTAGAAGTGGAGCTGGGCAGTGCAGAGCTTC
CTCTCAGGGAGCGGTGGGGCCCGAACTGCCACCAGGAGGTGGCCATGGCAGGGCAGGCTCCTGCCCTGT
GGACGGGGACGGACCGGTACGACTGCAGTGGCAGGAGCAGCCCTGGCCCAGGATGCTTTTCATAGGAAG
GCAAGTTCCCTGGACTCCGAGTTCCTATCGCTCCACCTCCTCGCCAGGCCTGTTCTCTCCCTGGGTCCTG
TCTTGAATGAGTCTAGACCTGTCGTTTGT
TCAGGAGACT

GACACTGAAGAAGCTTAAAATCACTTCACACAACAAAGTAGCACAAAGCAGTTTAAACAGAAAGAGCACAT
TTGTGGACTTCCAGATGGTCAGGAGATGAGCAAAGGATTGGTATGTGACTCTGATGCCCCAGCACAGTTA
CCCCAGCGAGCAGAGTGAAGAAGATGTTTGTGTGGGTTTTGTTAGTCTGGATTTCGGATGTATAAGGTGTG
CCTTGACTGTCTGATTTACTACACAGAGAACTTTTTTTTTTTTTTAAGATATATGACTAAAATGGACA
ATTGTTTACAAGGCTTAACTAATTTATTTGCTTTTTTAACTTGAACCTTTTCGTATAATAGATACGTTCT
TTGGATTATGATTTTAAGAAATTATTAATTTATGAAATGATAGGTAAGGAGAAGCTGGATTATCTCCTGT
TGAGAGCAAGAGATTTCGTTTTGACATAGAGTGAATGCATTTTCCCCTCTCCTCCTCCTGCTACCATTAT
ATTTTGGGGTTATGTTTTGCTTCTTTAAGATAGAAATCCCAGTTCTCTAATTTGGTTTTCTTCTTTGGGA
AACCAACATACAAATGAATCAGTATCAATTAGGGCCTGGGGTAGAGAGACAGAACTTGAGAGAAGAGA
AGTTAGTGATTCCCCTCTCTTTCTAGTTTGGTAGGAATCACCCCTGAAGACCTAGTCCTCAATTTAATTGTG
TGGGTTTTTTAATTTTCTAGAAATGAAGTGACTGAAACAATGAGAAAGAATACAGCACAAACCCTGAACAA
AATGTATTTAGAAATATATTTAGTTTATAGCAGAAGCAGCTCAATTGTTTGGTTGGAAAGTAGGGGAAA
TTGAAGTTGTAGTCACTGTCTGAGAATGGCTATGAAGCGTCATTTACATTTTTACCCCACTGACCTGCA
TGCCAGGACACAAGTAAAACATTTGTGAGATAGTGGTGGTAAGTGATGCACTCGTGTTAAGTCAAAGGC
TATAAGAAACACTGTGAAAAGTTCATATTCATCCATTGTGATTCTTTCCCCACGTCTTGCATGTATTACT
GGATTCCCACAGTAATATAGACTGTGCATGGTGTGTATATTTTCATTGCGATTTCTCTGTTAAGATGAGTTT
GTACTCAGAATTGACCAATTCAGGAGGTGTAAAAATAAACAGTGTTCTCTCTCTACCCCAAAGCCACTA
CTGACCAAGGTCTCTTCAGTGCCTCGCTCCCCTCTCTGGCTAAGGCATGCATTAGCCACTACACAAGTCA
TTAGTGAAAGTGGTCTTTTATGTCTCCCAGCAGACAGACATCAAGGATGAGTTAACCAGGAGACTACTC
CTGTGACTGTGGAGCTCTGGAAGGCTTGGTGGGAGTGAATTTGCCACACCTTACAATTGTGGCAGGATC
CAGAAGAGCCTGTCTTTTATATCCATTCCCTTGATGTCATTGGCCTCTCCCACCGATTTTCATTACGGTGC
CACGCAGTCATGGATCTGGGTAGTCCGGAACAAAGGAGGGAAGACAGCCTGGTAATGAATAAGATCC
TTACCACAGTTTTCTCATGGGAATACATAATAAACCTTTTCATCTTTTTTTTTTTTCTTTAAGAATTAA
AACTGGGAAATAGAAACATGAACGTGAAAAGTCTTGCAATGACAAGAGGTTTCATGGTCTTAAAAAGATAC
TTTATATGGTTGAAGATGAAATCATTCCTAAATTAACCTTTTTTTTTTAAAAAAAACAATGTATATTATGT
TCCTGTGTGTTGAATTTAAAAAAAATACTTTACTTGGATATTCATGTAATATATAAAGGTTTGGTG
AAATGAACTTTAGTTAGGAAAAAGCTGGCATCAGCTTTTCATCTGTGTAAGTTGACACCAATGTGTCATAA
TATTCTTTATTTTGGGAAATTAGTGTATTTTATAAAAATTTTAAAAAGAAAAAGACTACTACAGGTTAA
GATAATTTTTTTTACCTGTCTTTTCTCCATATTTTAAAGCTATGTGATTGAAGTACCTCTGTTTCATAGTTTC
CTGGTATAAAGTTGGTTAAAATTTTCATCTGTTAATAGATCATTTAGGTAATATAATGTATGGGTTTTCTAT
TGGTTTTTTTGCAGACAGTAGAGGGAGATTTTGTAAACAAGGGCTTGTTACACAGTGATATGGTAATGATAA
AATTGCAATTTATCACTCCTTTTCATGTTAATAATTTGAGGACTGGATAAAAGGTTTCAAGATTAAAAAT
TGATGTTCAAACCTTTGT

Figure 80: Domain Analysis of Human PRT3

Domain Name	begin	end	E-value
<u>RING</u>	12	52	1.06e-08
<u>SH3</u>	137	192	2.76e-19
<u>SH3</u>	199	258	4.84e-15
<u>low complexity</u>	366	384	-
<u>low complexity</u>	390	434	-
<u>SH3</u>	448	505	2.40e-19
<u>low complexity</u>	547	563	-
<u>low complexity</u>	652	668	-
<u>low complexity</u>	705	729	-
<u>SH3</u>	832	888	1.47e-14

Figure 81: Diagram of Human PRT3 Nucleic Acids

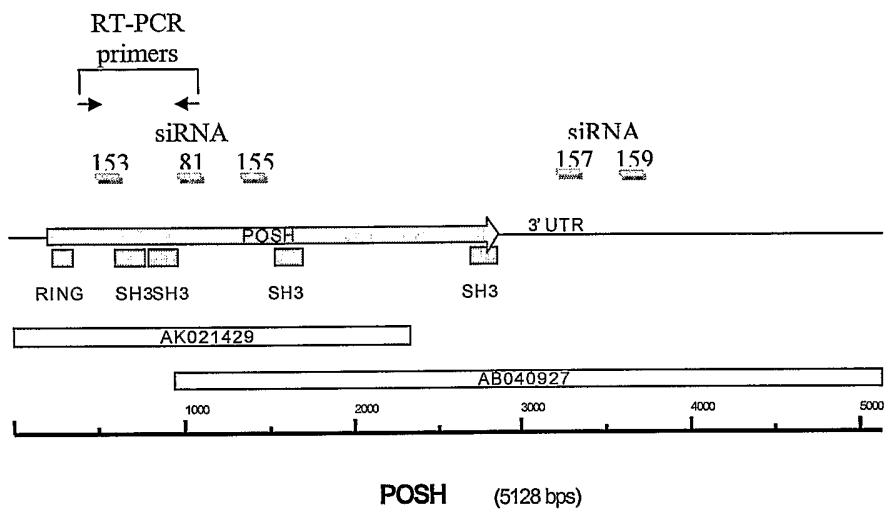


Figure 82: Reduction in Full Length PRT3 mRNA by siRNA Duplexes

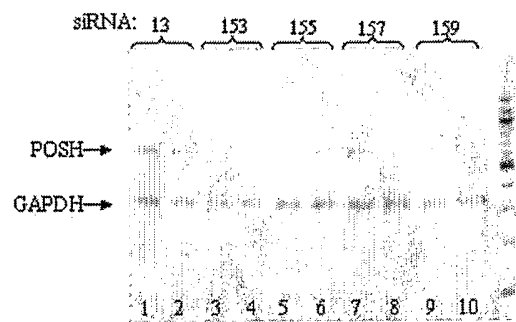


Figure 83: PRT3 Affects Release of VLP from Cells

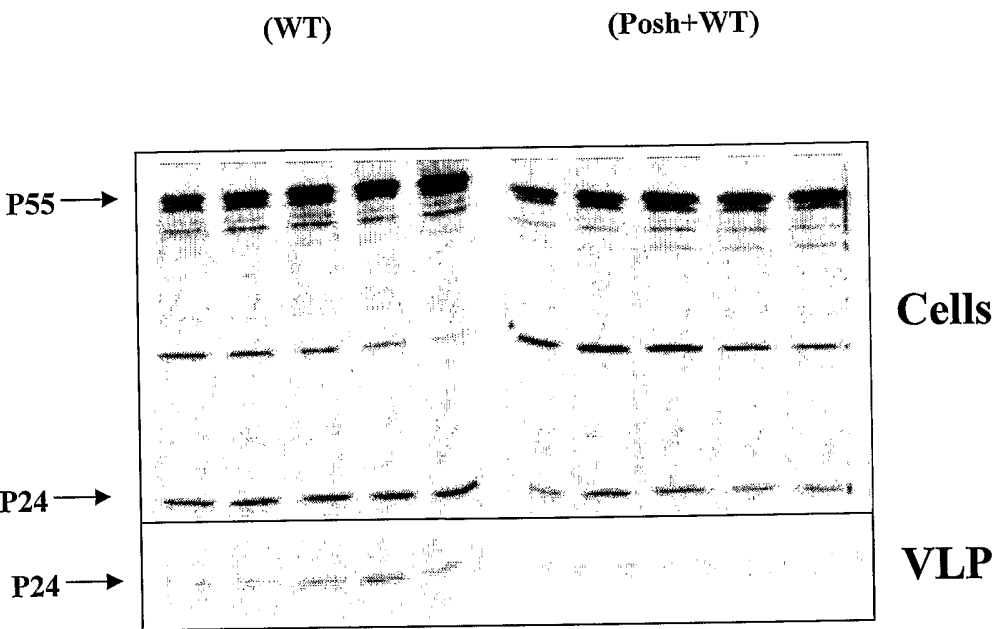


Figure 84: Release of VLP from Cells at Steady State

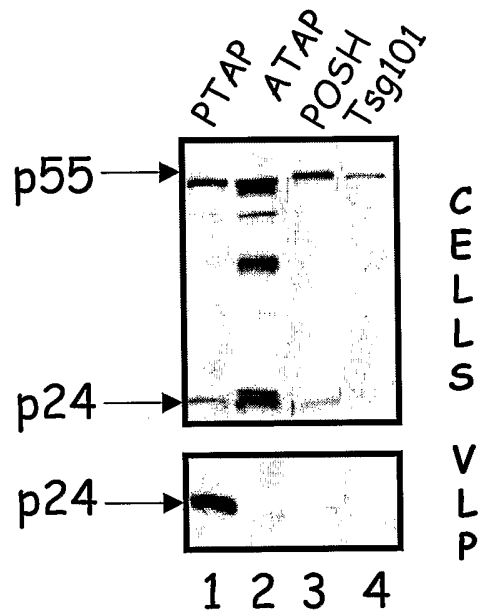


Figure 85: Exemplary HIV-1 Gag Nucleic Acid Sequence (Acc. No. NC_001802) (SEQ ID NO:80)

```
atggg tgcgagagcg tcagtattaa
gcgggggaga' attagatcga tgggaaaaaa ttcgggtaag gccaggggga aagaaaaaat
ataaattaaa acatatagta tgggcaagca gggagctaga acgattcgca gttaatcctg
gcctgttaga aacatcagaa ggctgtagac aaatactggg acagctacaa ccatcccttc
agacaggatc agaagaactt agatcattat ataatacagt agcaaccctc tattgtgtgc
atcaaaggat agagataaaa gacaccaagg aagctttaga caagatagag gaagagcaaa
acaaaagtaa gaaaaaagca cagcaagcag cagctgacac aggacacagc aatcagggtca
gccaaaatta ccctatagtg cagaacatcc aggggcaaat ggtacatcag gccatatcac
ctagaacttt aaatgcatgg gtaaaagtag tagaagagaa ggctttcagc ccagaagtga
taccatggtt ttcagcatta tcagaaggag ccaccccaca agatttaaac accatgctaa
acacagtggg gggacatcaa gcagccatgc aaatgttaaa agagaccatc aatgaggaag
ctgcagaatg ggatagagtg catccagtgc atgcagggcc tattgcacca ggccagatga
gagaaccaag gggaagtgac atagcaggaa ctactagtac ccttcaggaa caaataggat
ggatgacaaa taatccacct atcccagtag gagaaattta taaaagatgg ataatcctgg
gattaaataa aatagtaaga atgtatagcc ctaccagcat tctggacata agacaaggac
caaaggaacc ctttagagac tatgtagacc ggttctataa aactctaaga gccgagcaag
cttcacagga ggtaaaaaat tggatgacag aaaccttggt ggtccaaaat gcgaaccag
attgtaagac tattttaaaa gcattgggac cagcggctac actagaagaa atgatgacag
catgtcaggg agtaggagga cccggccata aggcaagagt tttggctgaa gcaatgagcc
aagtaacaaa ttcagctacc ataatgatgc agagaggcaa ttttaggaac caaagaaaga
ttgttaagtg tttcaattgt ggcaaagaag ggcacacagc cagaaattgc agggccccta
ggaaaaaggg ctgttggaag tgtggaaagg aaggacacca aatgaaagat tgtactgaga
gacaggctaa ttttttaggg aagatctggc cttcctacaa gggaaggcca gggaattttc
ttcagagcag accagagcca acagcccac cagaagagag cttcagggtc ggggtagaga
caacaactcc ccctcagaag caggagccga tagacaagga actgtatcct ttaacttccc
tcaggtcact ctttggcaac gacccctcgt cacaataa
```

Figure 86: Exemplary HIV-1 Gag Amino Acid Sequence (Acc. No. NP_057850) (SEQ ID NO:81)

```
1 mgarasvlsq geldrwekir lrpqgkkyk lkhivwasre lerfavnpql  
letsegcrqi  
61 lgqlqpslqt gseelrslyn tvatlycvhq rieikdtkea ldkieeqnk  
skkkaqqaaa  
121 dtghsnqvsq nypivqniqq qmvhqaispr tlnawvkve ekafspevip  
mfsalsegat  
181 pqdlntmlnt vggghqaamqm lketineeaa ewdrvhpvha gpiapgqre  
prgsdiagtt  
241 stlqeqigwm tnnppipvge iykrwiilgl nkivrmyspt sildirggpk  
epfrdyvdrf  
301 yktlraeqas qevknwmtet llvqnanpdc ktilkalgpa atleemmtac  
qgvvggpgghka  
361 rvlaeamsqv tnsatimmqr gnfrnqrkiv kcfncgkegh tarncraprk  
kgcwkcqkeg  
421 hqmkdcterq anflgkiwps ykgrpgnflq srpeptappe esfrsgvett  
tppqkqepid  
481 kelypltslr slfgndpssq
```

Figure 87: Exemplary HIV-1 p6 Amino Acid Sequence (SEQ ID NO:82)

1 lq srpeptappe esfrsgvett tppqkqepid kelypltslr slfgndpssq 51

Figure 88**Representative consensus terms for domains:****1. RING FINGER** - Consensus terms for the RING domain, SMART

016681/1-45 CKVCNENYSVDVE-----SHVPRVLT-CGH--SICQSCAAKQMS-
 CONSENSUS/80% C.lChp.....hhh..CtH...hp..Ch.thh..
 CONSENSUS/65% CslChp.....ptshhs.CsHh..aChtClpphht..
 CONSENSUS/50% CsIChp.....hpsshhs.CuHt..FCpsClpphht

016681/1-45 -----NSLILCKTC
 CONSENSUS/80% tCPhC
 CONSENSUS/65% pCPhC
 CONSENSUS/50% t.....hpCPlC

2. HECT - Consensus terms for the HECT domain, SMART (Pfam: PF00632)

CEC34D4_9 LDQAIRLMKHADRKAIVLEIEY-----TNEEGTGLGPTLEFYALVAAE
 CONSENSUS/80% ..p..plpt.....l.l.h.....sptshhsG..cEah..lspp
 CONSENSUS/65% hpspcL+pp.....lclpa.....sEpulDtGGhs+Eaa..llscp
 CONSENSUS/50% ppscc-LK+p.....LcicF.....pGEEGLDuGGlo+EFapLlo+E

CEC34D4_9 LQRK-----SLALWVCCCCDTHASKSGXKPIGYVRRVGGFLFPAPLPPGTDETKRAAD--
 CONSENSUS/80% hhps.....shh.ht.tp.....hh..ss.s...t..p.....
 CONSENSUS/65% hhss.....hshaphssssp.....hhh.ssts.h.sttch....
 CONSENSUS/50% lFNP.....shuLFcassscp.....tLassPasushtssc+lt....
 CEC34D4_9 MFRVLGVFLAKVLLD-----GRLVDLPLSRPFLKLLVHPQIGXILSLDDFEVNPVKG
 CONSENSUS/80% hathlGhhhuhs1hp.....tphl-h.hs..hhp.lhth.....phpDhtph-s.hh
 CONSENSUS/65% hFchlGp1lGhs1hc.....ppl1DhhFsp.hhKh1lph.....hshpDhphDs.ha
 CONSENSUS/50% aF+FlGR1lGhAlac.....scL1DlpFs+sFYKpLLscs....ls1cDLcs1DPshY

CEC34D4_9 --SFLKELRALAQRKRLIENDTSXELKLHIKGSTCRVEDLALNFTVNPPSKVFQYAEML
 CONSENSUS/80% ..psl..l.....hp...tt...p..h.h....t..t.h..hcl
 CONSENSUS/65% ..psLhhl.....hptshsp..hs..lsF.hs.pp.hstst.h-l
 CONSENSUS/50% ..cSLptL.....lcscssch.h-.LsF.lshasphGpsss1EL

CEC34D4_9 VDGGSDIDVTIDNVEQYVEKCEEY-YLNTGIAYQMRAFRDGFDRVFPLRTLRAVSPEEVQ
 CONSENSUS/80% h.sGpph.VT.pNh.palpphhph.hh.tthtp.hpuhhpGh.plhs.phht.hhstp-l
 CONSENSUS/65% hssGps1.VTppNhtEYlphhsa.hhtptlpcphpAFncGh.plls.phlp.hFetp-l
 CONSENSUS/50% pPGGcsIsVtpcn+cEYVcLls-a.+Ls+ulccQhcAF+cGFscVlPpchLp.lFssCEL

CEC34D4_9 RLLSGEQCPWSRDDILNYTEPKLGYTRESFGF----LRFVDVMEALTAQERKNFLQFAT
 CONSENSUS/80% phhhsG...phshpshhp.s.ht.thtsp..l...phFWphlpphs.pc+hthl.FhT
 CONSENSUS/65% c1llsGht.plshp-hcppsap.tYstsp.hl...phFWcllcpshpcc+tpHlpFlT
 CONSENSUS/50% -LLluGss.-IDlcDL+ssTpYcsGYocsspsI....+WFWcll-sFssEERp+LLQFVT

CEC34D4_9 GCSSLP--PGGLANLHP-----RLTIVRKVE---SGDGSYPSVNTCVHYLKLPEYSSSA
 CONSENSUS/80% ussphp...tGhttl.....phtl.....ttp.LPpupTChs.L.LP.Ypohp
 CONSENSUS/65% GosplP..htGftsL.s.....phsIttts.....spspLPoupTChNhLpLp.YsSpp
 CONSENSUS/50% GoSRlP..lsGFusLps.....+FsIcpast.....cscLPoAatCaNpLcLpYSSKc

CEC34D4_9 ILRERLLTAINE-----KGFHLN
 CONSENSUS/80% hlcp+LhhAlpp.....tsFth.
 CONSENSUS/65% hl+p+LhhAipp.....tpuFuht
 CONSENSUS/50% pLRcKLhA1cs.....scGFGLs

3. RCC1 - Consensus terms for the RCC1 domain, Pfam PF00415

RCC_HUMAN/34-84 PGLVLTGQ.GDVGQLGLGENVMERK.....KPALV.....SIPEDVVQAEAGGMHTVCLSK P18754
 SS CBEEEEEE..CTTSTTCSCCTTCCEEE.....EEEE.....CCSSCEEEEEECSSEEEEET SS
 SA 303000001..0530000216923334.....40140.....817350000200030000005 SA
 RCC_HUMAN/85-136 SQGVYSFGC..NDEGALGRDTSVEGSE.....MVPQKV.....ELQEKVVQVVSAGDSHTAALTD P18754
 SS TSCEEEEEEC..CTTSTTCSCCCTTGG.....GSCCEC.....CCSCCEEEEEECSSEEEEET SS
 SA 502000002..0710000172778112.....042140.....629340100010310000001 SA
 RCC_HUMAN/137-189 DGRVFLWGSFRDMNGVIGLEPMKKS.....VPVQV.....QLDVPVVKVASGNDHLVMLTA P18754
 SS TSCEEEEEECESSEESSBETBCEE.....EEEC.....CCSSCEEEEEECSSEEEEET SS
 SA 102010000031771200025265414.....61340.....728230000110200000000 SA
 RCC_HUMAN/363-416 DGRVFAWGM..GTNYQLGTGQDEDAN.....SPVEM...MGKQLENRVVLSVSSGGQHTVLLVK P18754
 SS TSCEEEEEEC..CTTSTTCSSSCSCEE.....EEEC...CSTTTTEEEEEEECSSEEEEEEE SS
 SA 312000001..051100031546312.....00330...6183167110220001230000001 SA
 RCC_HUMAN/312-362 EGKAYSLGR..AEYGRLLGEGAEBSK.....IPTLI.....SRLPAVSSVACGASVGYAVTK P18754

```
SS          TSCEEEEC..CGGGTTCCTTCCEE.....EEEC.....CSSSSEEEECSSSEEEET SS
SA          10000001..0440000106717433.....41210.....00002042101020000006 SA
RCC_HUMAN/190-257 DGDLYTLGC..GEQQLGRVPELAFANRGGGRQLERLLVPKCVMLKSRGSRGHVRFQDAFCGAYFTFAISH P18754
SS          TSCEEEEC..CTTSTTSCCGGCTTTTGGGCHHHHSCERCCBCSSCCSBCEEEETEEEEET SS
SA          500010000..01200001036512731375205200414106061687752130000110020000003 SA
RCC_HUMAN/258-310 EGHVYGFGL..SNYHQLGTPGTESCF.....IPQNL...TSFKNSTKSWVGFSGGQHHTVCM P18754
SS          TCCEEEEC..CTTSTTSCSSCCEE.....EEEC...GGGCTTCCEEEECSSSEEEET SS
SA          412000000..032200008447211.....60320....720535110031001031000000 SA
SS_cons     TSCEEEECCTTSTTSCCCSCCEETGGGCHHHHEEECCBCSCCCSSCEEEECSSSEEEET SA_cons
SA_cons     3020000010305310001346542331375205200412306064563535230110110120000002 SA_cons
```

4. WW domain - Consensus terms for the WW domain, SMART (Pfam: PF00397)

```
054971/1-33      PLPPGWEKRTDSN-GRVYFVN---HNTRITQWEDPRS
CONSENSUS/80%    .h..sW..hhs.p.sh.aahs....stpopWptPt.
CONSENSUS/65%    slssWppthsss.GpHYhs...ppocpopWpcPp.
CONSENSUS/50%    sLPsGwccttsss.G+sYYaN...ppT+copW-cPss
```

5. C2 - Consensus terms for the C2 domain, SMART (Pfam: PF00168)

```
JC5772/1-104     VTQVHVLGAAGLKDSPT-----
CONSENSUS/80%    .h.l.lhpsptl.....
CONSENSUS/65%    hlplplhpupsl.ths.....
CONSENSUS/50%    pLplplpAcslsspch.....
```

```
JC5772/1-104     -GANSYVIKCEGD-----
CONSENSUS/80%    ...ssas.h.h.....
CONSENSUS/65%    .hs-sYlplphs.....
CONSENSUS/50%    tpsDPYVclplsattt.....
```

```
JC5772/1-104     -----KVRSAVQK
CONSENSUS/80%    .....pTphh
CONSENSUS/65%    .....hh+Tphh
CONSENSUS/50%    .....pt+Tclh
```

```
JC5772/1-104     GTS-----TPEYNV---KGIFYRKKL-----SQPITVQ
CONSENSUS/80%    tps.....sP.app...th.h.h.....t.l.ht
CONSENSUS/65%    pps.....sPhWs...ph.hph.....tphlplp
CONSENSUS/50%    cssh.....NPpWse...sFpFslsst.....sspLplp
```

```
JC5772/1-104     VWNHRVL-----KDEFLGQVHLKADP-----DN
CONSENSUS/80%    lhs.p.....pp.lGth.h.....ht
CONSENSUS/65%    Vh-psth.....tsphlGpstls.....lp
CONSENSUS/50%    VaDc-pht.....pcchlGpspls.....ls
JC5772/1-104     LQALHTLHLRD
CONSENSUS/80%    ph..t.....
CONSENSUS/65%    plhttttt.hh
CONSENSUS/50%    cltssppstth
```

6. SH2 - Consensus terms for the SH2 domain, SMART (Pfam: PF00017)

```
SRC_AVIS2/1-91    EEWYFGKI----TRRESERLLLNP-----ENPRGTFVLVRESET-----
CONSENSUS/80%    ..aaas.h....sptpupphL.t.....GsallR.opp.....
CONSENSUS/65%    t.WaaGtl....sRppAcphLpp.....pGsFLlRpSps.....
CONSENSUS/50%    psWaHGsl....oRp-AEcLlps.....ttcGsFLlRcScs.....
```

```
SRC_AVIS2/1-91    -----TKGAYCLSVSDFDN-----
CONSENSUS/80%    .....tthslorrh.t.....
CONSENSUS/65%    .....spasLslhph.....
CONSENSUS/50%    .....psusasLslppp.....
```

```
SRC_AVIS2/1-91    -----AKGLNVKHYKIRKLD-----
CONSENSUS/80%    .....tthhchhl...t.....
CONSENSUS/65%    .....spspHhplppps.....
CONSENSUS/50%    .....spl+HapIppsst.....
```

```
SRC_AVIS2/1-91    -----GGFYITS-----RTQFSS--LQQLVAYYSKHADG
CONSENSUS/80%    ....t.h.htt.....a.s.l.pLlpha.tt...
CONSENSUS/65%    ....stahlst.....htFso.l.cLlpaappshs
CONSENSUS/50%    ....spatlss.....tppFso..Ls-Llsayppssts
```

7. transmem - Transmembrane region prediction, SWISSPROT

SWISSPROT make use of a program that is based on four different methods:

1. Eisenberg, Schwarz, Komaromy and Wall (J. Mol. Biol. 179:125-142(1984));
2. Rao and Argos (Eur. J. Biochem. 128:565-575(1982));
3. Klein, Kanehisa and DeLisi (Biochim. Biophys. Acta 815:468-476(1985));
4. MEMSAT (Jones, Taylor and Thornton) (Biochemistry 33:3038-3049(1994)).

When methods 1 and 4 show the protein to have hydrophobic regions, the transmembrane regions are added according to the results of method 1. The result is:

FT TRANSMEM x y POTENTIAL.

Figure 89



Comparative Sequence Analysis



SMART Simple Modular Architecture Research Tool

Amino Acid Grouping

Below you see the grouping of amino acids to classes as well as the classes abbreviation (the key) used within the consensus sequence

Class	Key	Residues
alcohol	o	S, T
aliphatic	l	I, L, V
any	.	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
aromatic	a	F, H, W, Y
charged	c	D, E, H, K, R
hydrophobic	h	A, C, F, G, H, I, K, L, M, R, T, V, W, Y
negative	-	D, E
polar	p	C, D, E, H, K, N, Q, R, S, T
positive	+	H, K, R
small	s	A, C, D, G, N, P, S, T, V
tiny	u	A, G, S
turnlike	t	A, C, D, E, G, H, K, N, Q, R, S, T

If you want to change the grouping and/or keys or if you want the consensus sequence calculated using different thresholds, please use our [Consensus Server](#)



Last modified: Feb. 16, 1998