PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C07K 14/415, C12N 15/63, C07K 16/18, C12N 15/11, G01N 33/53

A2

(11) International Publication Number: **WO 00/40607**

(43) International Publication Date: 13 July 2000 (13.07.00)

(21) International Application Number: PCT/US00/00459

(22) International Filing Date: 6 January 2000 (06.01.00)

(30) Priority Data:
09/227,278 6 January 1999 (06.01.99) US

(71) Applicant *(for all designated States except US)*: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: DUHL, David [US/US]; 2720 Camino Lenada, Oakland, CA 94611 (US).

(74) Agents: BLACKBURN, Robert, P. et al.; Chiron Corporation, 4560 Horton Street, Emeryville, Ca 94608–2916 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, 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, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

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

(54) Title: DNA ENCODING T–CELL LYMPHOMA INVASION AND METASTASIS 2 AND A PROCESS FOR ITS USE

(57) Abstract

The present invention provides nucleic acid sequences encoding TIAM2 proteins. The invention also includes diagnostic assays, expression vectors, antisense molecules, ribozymes, and host cells to express the polypeptide encoded by the nucleic acid sequence. The present invention also includes claims to the polypeptide sequence encoded by the nucleic acid sequences.

FOR THE PURPOSES OF INFORMATION ONLY

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

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

DNA ENCODING T-CELL LYMPHOMA INVASION AND METASTASIS 2 AND A PROCESS FOR ITS USE

FIELD OF THE INVENTION

The present invention is in the field of molecular biology. More specifically, the present invention relates to a polynucleotide sequence and corresponding TIAM2 protein.

BACKGROUND OF THE INVENTION

The TIAM1 gene and characteristics of the encoded protein are discussed in Habets et al., *Cell 77*:537-549 (1994); Hart, M.J. et al., *J. Biol. Chem. 269*:62-65 (1994); Haslam et al., *Nature 363*:309-310 (1993); and Mayer, B.J. et al., *Cell 73*:629-630 (1993).

SUMMARY

The invention relates to a polynucleotide comprising a polynucleotide of SEQ ID NO:8.

The invention also relates to a polypeptide encoded by SEQ ID NO:8.

The invention further relates to a polynucleotide having at least 80% sequence identity to SEQ ID NO:8.

The invention still further relates to a polynucleotide having at least 85% sequence identity to SEQ ID NO:8.

The invention also relates to a polynucleotide having at least 90% sequence identity to SEQ ID NO:8.

The invention further relates to a polynucleotide having at least 95% sequence identity to SEQ ID NO:8.

The invention relates to a polynucleotide comprising a polynucleotide of SEQ ID NO:9.

The invention also relates to a polypeptide encoded by SEQ ID NO:9.

The invention further relates to a polynucleotide having at least 80% sequence identity to SEQ ID NO:9.

The invention still further relates to a polynucleotide having at least 85% sequence identity to SEQ ID NO:9.

The invention also relates to a polynucleotide having at least 90% sequence identity to SEQ ID NO:9.

The invention further relates to a polynucleotide having at least 95% sequence identity to SEQ ID NO:9.

The invention relates to a polynucleotide comprising a polynucleotide of SEQ ID NO:10.

The invention also relates to a polypeptide encoded by SEQ ID NO:10.

The invention further relates to a polynucleotide having at least 80% sequence identity to SEQ ID NO:10.

The invention still further relates to a polynucleotide having at least 85% sequence identity to SEQ ID NO:10.

The invention also relates to a polynucleotide having at least 90% sequence identity to SEQ ID NO:10.

The invention further relates to a polynucleotide having at least 95% sequence identity to SEQ ID NO:10.

The invention relates to a polypeptide having at least 80% sequence identity to the amino acid sequence of SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13.

The invention also relates to a polypeptide having at least 85% sequence identity to the amino acid sequence of SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13.

The invention further relates to a polypeptide having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13.

The invention still further relates to a polypeptide having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13.

The invention further relates to an antibody capable of specifically binding to a polypeptide encoded by SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10, or by a polynucleotide having at least 80% sequence identity to SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

The invention further relates to an antibody capable of specifically binding to a polypeptide having the amino acid sequence of SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13.

The invention also relates to a method of detecting expression of a TIAM2 expression product in a biological sample, comprising obtaining polypeptides from the biological sample, contacting the polypeptides with an antibody capable of specifically binding to TIAM2, and detecting the presence or absence of antibody-polypeptide complexes.

The invention also relates to a method of obtaining a TIAM2 polypeptide, comprising expressing a polynucleotide having at least 80% sequence identity to SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

The invention further relates to a vector comprising a polynucleotide having at least 80% sequence identity to SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

The invention still further relates to a vector comprising a polynucleotide having at least 80% sequence identity to SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10, and at least one regulatory region.

The invention also relates to a host cell comprising a polynucleotide having at least 80% sequence identity to SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

The invention further relates to a method of detecting a deletion of a region of chromosome 6q25 in ovarian cancer, comprising comparing the polynucleotide sequence of region of chromosome 6q25 of an ovarian tissue sample

suspected of being cancerous with the polynucleotide sequence of SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10 and determining the presence or absence of the sequence in the polynucleotides of the tissue sample.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Human brain-(A) and tissue-specific (B) blots containing 2 μg/lane of poly(A)$^+$ mRNA were hybridized with TIAM-specific probes that recognize both the long ($TIAM_L$) and the short ($TIAM_S$) forms of TIAM2. A probe derived from the 5' end of the 4.4-Kb message hybridized to only the long form (C), while a 3'-specific probe hybridized to both the long and the short forms (D), on the brain-specific Northerns.

Figure 2 shows expression of TIAM2 and mRNA in mouse brain. (A) Whole mount of a E13.5 embryo brain. TIAM2 transcripts are expressed throughout the developing telencephalon (T). (B) Parasagittal section through an E13.5 telencephalon demonstrating that TIAM2 mRNA is localized to cells near the pial (P) surface; labeling is not seen near the ventricular surface (V). (C) Coronal section through the adult mouse forebrain. TIAM2 transcripts are localized to the cerebral cortex (Ctx) and caudate putamen (CP), with intense labeling in the ependyma (small arrow) and in the indusium griseum (arrow). (D) Higher magnification view of labeling in the emendyma (arrowhead) lining the lateral ventricle (LV) below the corpus callosum. (CC). (E) TIAM2 transcripts in the olfactory bulb, glomerular layer (arrowhead), mitral cell layer (arrow). (F) Coronal section demonstrating labeling in the thalamus (arrowhead) and fasciola cinerea (arrow). (G) Hippocampal expression of TIAM2 shows expression in the granule cells of the dentate gyrus (DG) and the pyramidal cells of CA2.

Figure 3 shows a comparison of the carboxy-terminus amino acid sequence of TIAM1 protein (amino acids 576-1591) with the amino acid sequence of $TIAM2_L$. The EX domain, alternatively splice region, DHR domain, DH domain, and PH domain are underlined, and the initiator methionine for $TIAM2_S$ (amino acid 452) is in boldface and designated with an arrow.

Figure 4 shows the results of a GDP-GTP exchange assay. Partially purified His-TIAM2$_S$ protein was assayed for the ability to stimulate GDP-GTP exchange activity with immunoprecipitated Rac (A) or Ras (B) proteins. (A) The average of three separate reactions in which Rac was incubated with [^{32}P]GTP in the presence of His-TIAM2$_S$ (◇), SOS (○), or buffer alone (□). (B) A representative experiment in which Ras was incubated with [^{32}P]GTP in the presence of TIAM2$_S$ (●), SOS (▲), or buffer alone (X).

Figure 5 shows the expression of TIAM2 mRNA in tumor tissue and normal tissues.

DETAILED DESCRIPTION OF THE INVENTION

TIAM2 (DP-75), for T-cell lymphoma invasion and metastasis 2, relates to a novel DNA and amino acid sequence that has some sequence homology with TIAM1. (Habets *et al.*, *Cell 77*:537-749 (1994), and Habets *et al.*, *Oncogene 10*:1371-1376 (1995)). Overexpression of full length or truncated forms of TIAM1 increases the metastatic potential of lymphoma cells in mice. TIAM1 is a member of a family GDP dissociation stimulators (GDSs) which are proteins that activate Rho-like and Rac-like GTPases. GDSs as well as Rho and Rac have oncogenic potential.

According to the invention, a novel protein, TIAM2, with high homology to TIAM1, has been identified. TIAM2 is expressed as ~4.4- and ~3.3-kb messages that encode long (TIAM2$_L$) and short forms (TIAM2$_S$) of the TIAM2 protein. TIAM2$_L$ begins in a region with identity to the coiled-coil region of TIAM1 and includes both the EX and the PDZ domains found within TIAM1. Both TIAM2$_L$ and TIAM2$_S$ contain regions of identity to the DH and carboxyl-terminal PH domains of TIAM1. An additional level of complexity is added by alternative splicing of the ~4.4-kb message, which leads to the insertion of 24 amino acids between the EX and the DHR domains.

While analysis of the ~4.4-kb message identified an ATG at nucleotide 51 that is in good context for translation, further analysis revealed that the ORF can be extended 5' of the ATG through the first 51 nucleotides. This extended ORF could

indicate that there are additional 5' sequences that would extend $TIAM2_L$. Initiation likely begins at nucleotide 51 since the size of the clone cDNA (4586 nt) equals or exceeds the estimated size of the cerebellum message (~4.4-kb) and because, despite numerous attempts, additional 5' sequences were not identified.

A His-tagged version of $TIAM2_S$ was expressed in a baculovirus system, which is the preferred system for TIAM2 expression, and purified on a nickel column. His-$TIAM2_S$ was compared to SOS and was shown to have GEF activity that prefers RAC over RAS.

The high similarity between TIAM1 and TIAM2 protein coding sequences, expression patterns, and GEF activity suggest that, like TIAM1 (Habets et al., *Cell 77*:537-549, 1994; Hordijk et al., *Science 278*:1464-1466, 1997; Ehler et al., *Mol. Cell. Neurosci. 9*:1-12, 1997; van Leeuwen et al., *J. Cell Biol., 139*:797-807, 1997), TIAM2 may play a role in neural development. The high level of TIAM2 expression throughout the E13.5 telencephalon suggests that TIAM2 may play an important role in development of the forebrain. At E13.5 postmitotic neurons are being produced; these cells must migrate to their correct positions within the developing brain and send out growth cones to establish contacts with their correct targets. These processes all require coordinating actin-based cytoskeletal rearrangements with environmental cues. There is strong evidence that the Rac/cdc42 GTPases play important roles in neuritogenesis and target recognition within the nervous system (Albertinazzi et al., *J. Cell biol. 142*:815-825, 1998; Hing et al., *Cell 97*:853-863, 1999; van Leewen et al., *supra*, 1997). Further analysis of TIAM2 expression during mouse embryogenesis may reveal a correlation with specific morphogenetic events or with specific GTP-ase family members, regulators, or effectors. It is likely that some of the specificity established during neural development is due to cell-type-specific expression of GTPase signaling pathway components such as TIAM1, RAC3, and TIAM2.

Unlike TIAM1, TIAM2 continues to be expressed at high levels in the adult forebrain. The areas of high expression in the adult are regions of synaptic plasticity (cortex, hippocampus) or neurogenesis (dentate gyrus, ependyma/subependyma) in the adult (Gould et al., *Trends Cogn. Sci. 3*:186-192, 1999;

Temple, *Curr. Biol. 9*:R397-399, 1999; Lowenstein and Parent, *Science 283*:1126-1127, 1999).

TIAM2 may play a role in metastasis or tumor suppression. Radiation hybrid mapping of TIAM2 places it in a 4-cM region of chromosome 6q25 that is frequently deleted in ovarian cancer (Colitti et al., *Oncogene 16*:555-559, 1998). Since the majority of ovarian cancer is epithelial in nature, and since TIAM1 (21q22.1; Chen and Antonarakis, *Genomics 30*:123-127, 1995) has been suggested to act as a tumor suppressor in epithelial cells, this might suggest a role for TIAM2 in ovarian cancer.

The expression of alternative forms of TIAM2 suggests distinct roles for $TIAM2_L$ and $TIAM2_S$. Both the EX and the PDZ domains found in $TIAM2_L$ have been suggested to play roles in protein-protein interactions and intracellular localization of proteins. Since $TIAM2_S$ lacks both the EX and the PDZ domains of $TIAM2_L$, the two forms may be localized to different regions of the cell. Understanding the alternate forms of TIAM2 extends knowledge of the complexity of this emerging family of GEFs and may ultimately yield insight into the divergent roles of TIAM1 as both a promoter of invasiveness and a tumor suppressor.

GDSs are implicated in various cancers and, based on the discussion herein, TIAM2 may be useful to diagnose cancerous cells. Many techniques may be used to diagnose whether tissue samples possess TIAM2-containing tumor tissue. For example, reverse transcription and PCR amplification of the RNA of a tumor sample to identify the presence of TIAM2 mRNA sequences (see Sambrook, *et al.*, Molecular Cloning; A Laboratory Manual, Second Edition (1989), chapter 14 or Gaugler *et al.*, *J. Exp. Med 179*:921-930 (1994)). Also, immunohistochemical techniques or ELISA assays may be used to identify TIAM2-expressing tumors. For example, the TIAM2 protein can be recombinantly expressed and monoclonal antibodies can be prepared according to methods that are known in the art. For example, the methods shown in EP 174,204, Kohler and Milstein, *Nature 256*:495-497 (1975), Fong *et al.*, *J. Immun. Meth. 70*:83-90 (1984), GB 2,086,937, 2,113,715, EP 57,107, 62,409, EP 118,893, EP 124,301, and EP 131,878 are suited to the present invention. The anti-TIAM2

monoclonal antibodies can then be used in the standard assays recited above or those assays that are otherwise known in the art.

Monoclonal antibodies may also be used therapeutically. Anti-TIAM2 monoclonal antibodies can be administered by means known in the art. Preferably, the antibodies are administered parenterally or subcutaneously, more preferably, they are administered intravenously. The monoclonal antibodies can be administered in combination with other agents designed to promote the activity of the antibodies or to treat the underlying condition involving the TIAM2 expressing cell.

Additionally, branched DNA testing may be performed to assay for TIAM2 DNA as shown in U.S. Patent Nos. 5,124,246, and 4,868,105 (hereby incorporated by reference in their entireties). TIAM2 nucleic acid probe molecules for the branched DNA testing are preferably from 10 to 50 bases in length, more preferably, between 15 and 40 bases in length, most preferably, between 20 and 30 bases in length

Ribozymes may be designed to act on the TIAM2 sequence identified in SEQ ID NO:1, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof. (Kashani-Cabet and Scanlon, *Cancer Gene Therapy*, *2*:213-223 (1995), hereby incorporated by reference in its entirety.) To obtain cellular expression, the ribozyme gene is cloned into an available vector and transfected into the cells of choice. Different vectors may be chosen based on the target cell to be infected. For example, respiratory cells may be targeted by an adeno or adeno associated virus (AAV) vector. Appropriate promoters may be inserted into these vectors to ensure regulatable expression. (Kashani-Cabet at p. 216).

Antisense molecules can be developed based on the TIAM2 sequence shown in SEQ ID NO:1, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, or SEQ ID NO:10. For example, see U. S. Patent Nos. 5,491,133 and 5,271,941 which are hereby incorporated by reference in their entireties.

Antisense RNA sequences have been described as naturally occurring biological inhibitors of gene expression in both prokaryotes (Mizuno, T., Chou, M-Y, and Inouye, M. (1984), *Proc. Natl. Acad. Sci. USA 81*, (1966-1970)) and eukaryotes (Heywood, S. M. *Nucleic Acids Res., 14*, 6771-6772 (1986) and these sequences

function by hybridizing to complementary mRNA sequences, resulting in hybridization arrest of translation (Paterson, B. M., Roberts, B. E., and Kuff, E. L., (1977) *Proc. Natl. Acad. Sci. USA, 74*, 4370-4374.

Antisense oligodeoxynucleotides are short synthetic nucleotide sequences formulated to be complementary to a specific gene or RNA message. Through the binding of these oligomers to a target DNA or mRNA sequence, transcription or translation of the gene can be selectively blocked and the disease process generated by that gene can be halted. The cytoplasmic location of mRNA provides a target considered to be readily accessible to antisense oligodeoxynucleotides entering the cell; hence much of the work in the field has focused on RNA as a target.

Antisense therapy is the administration of oligonucleotides which bind to a target polynucleotide located within the cells. These oligonucleotides are usually exogenous, but they can be endogenously expressed. The term "antisense" refers to the fact that such oligonucleotides are complementary to their intracellular targets, *e.g.*, TIAM2. See for example, Jack Cohen, Oligodeoxynucleotides, Antisense Inhibitors of Gene Expression, CRC Press, 1989; and *Synthesis 1*:1-5 (1988).

The TIAM2 antisense oligonucleotides of the present invention include derivatives such as S-oligonucleotides (phosphorothioate derivatives or S-oligos, see, Jack Cohen, supra) which exhibit enhanced cancer cell growth inhibitory action. The TIAM2 antisense oligonucleotides of the present invention may be RNA or DNA which are complementary to and stably hybridize with the TIAM2 genome or the corresponding mRNA. Use of an oligonucleotide complementary to this region allows for the selective hybridization to TIAM2 mRNA and not to other mRNAs.

Preferably, the TIAM2 antisense oligonucleotides of the present invention are 15 to 40-mer fragments of the antisense DNA molecules which hybridize to TIAM2 mRNA. Alternatively, the preferred TIAM2 antisense oligonucleotide is a 20- to 30-mer oligonucleotide which is complementary to a region in TIAM2. Included in the present invention are pharmaceutical compositions comprising an effective amount of at least one of the TIAM2 antisense oligonucleotides of the invention in combination with a pharmaceutically acceptable carrier. In one embodiment, a single

TIAM2 antisense oligonucleotide is utilized. In another embodiment, two TIAM2 antisense oligonucleotides are utilized which are complementary to adjacent regions of the TIAM2 genome.

Administration of two TIAM2 antisense oligonucleotides which are complementary to adjacent regions of the TIAM2 genome or corresponding mRNA may allow for more efficient inhibition of TIAM2 genomic transcription or mRNA translation, resulting in more effective inhibition of cancer cell growth. Preferably, the TIAM2 antisense oligonucleotide is coadministered with an agent which enhances the uptake of the antisense molecule by the cells. For example, the TIAM2 antisense oligonucleotide may be combined with a lipophilic cationic compound which may be in the form of liposomes.

The use of liposomes to introduce nucleotides into cells is taught, for example, in U.S. Pat. Nos. 4,897,355 and 4,394,448, the disclosures of which are incorporated by reference in their entireties. See also U.S. Pat. Nos. 4,235,871, 4,231,877, 4,224,179, 4,753,788, 4,673,567, 4,247,411, 4,814,270 for general methods of preparing liposomes comprising biological materials.

Alternatively, the TIAM2 antisense oligonucleotide may be combined with a lipophilic carrier such as any one of a number of sterols including cholesterol, cholate and deoxycholic acid. A preferred sterol is cholesterol. In addition, the TIAM2 antisense oligonucleotide may be conjugated to a peptide that is ingested by cells. Examples of useful peptides include peptide hormones, antigens or antibodies, and peptide toxins. By choosing a peptide that is selectively taken up by the neoplastic cells, specific delivery of the antisense agent may be effected.

The TIAM2 antisense oligonucleotide may be covalently bound via the 5' H group by formation of an activated aminoalkyl derivative. The peptide of choice may then be covalently attached to the activated TIAM2 antisense oligonucleotide via an amino and sulfhydryl reactive hetero bifunctional reagent. The latter is bound to a cysteine residue present in the peptide. Upon exposure of cells to the TIAM2 antisense oligonucleotide bound to the peptide, the peptidyl antisense agent is endocytosed and

the TIAM2 antisense oligonucleotide binds to the target TIAM2 mRNA to inhibit translation. See PCT Application Publication No. PCT/US89/02363.1.

The TIAM2 antisense oligonucleotides and the pharmaceutical compositions of the present invention may be administered by any means that achieve their intended purpose. For example, administration of the antisense compounds or other compounds of the present invention may be by parenteral, subcutaneous, intravenous, intramuscular, intra-peritoneal, or transdermal routes.

The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired. Compositions within the scope of this invention include all compositions wherein the TIAM2 antisense oligonucleotide is contained in an amount which is effective to achieve inhibition of proliferation and/or stimulate differentiation of the subject cancer cells.

While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typically, the TIAM2 antisense oligonucleotide may be administered to mammals, *e.g.*, humans, at a dose of 0.005 to 1 mg/kg/day, or an equivalent amount of the pharmaceutically acceptable salt thereof, per day of the body weight of the mammal being treated.

In addition to administering the TIAM2 antisense oligonucleotides as a raw chemical in solution, the TIAM2 antisense oligonucleotides may be administered as part of a pharmaceutical preparation containing suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the TIAM2 antisense oligonucleotide into preparations which can be used pharmaceutically. The term pharmaceutically acceptable refers to compounds and compositions which may be administered to mammals without undue toxicity. Exemplary pharmaceutically acceptable salts include mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like.

Suitable formulations for parenteral administration include aqueous solutions of the TIAM2 antisense oligonucleotides in water-soluble form, for example,

water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran. Optionally, the suspension may also contain stabilizers.

The antisense oligonucleotides of the present invention may be prepared according to any of the methods that are well known to those of ordinary skill in the art. Preferably, the antisense oligonucleotides are prepared by solid phase synthesis. See, Goodchild, J., *Bioconjugate Chemistry*, *1*:165-167 (1990), for a review of the chemical synthesis of oligonucleotides. Alternatively, the antisense oligonucleotides can be obtained from a number of companies which specialize in the custom synthesis of oligonucleotides.

TIAM2 polypeptides, polynucleotides, or antibodies can be administered orally, topically, or by parenteral means. including subcutaneous and intramuscular injection, implantation of sustained release depots, intravenous injection, intranasal administration, and the like. When used to treat tumors, it may be advantageous to apply the TIAM2 polynucleotides or antibodies, for example, directly to the site, *e.g.*, during surgery to remove the bulk of the tumor. Accordingly, TIAM2 polypeptides, polynucleotides, or antibodies may be administered as a pharmaceutical composition comprising a pharmaceutically acceptable excipient. Such compositions may be aqueous solutions, emulsions, creams, ointments, suspensions, gels, liposomal suspensions, and the like. Suitable excipients include water, saline, Ringer's solution, dextrose solution, and solutions of ethanol, glucose, sucrose, dextran, mannose, mannitol, sorbitol, polyethylene glycol (PEG), phosphate, acetate, gelatin, collagen, Carbopol®, vegetable oils, and the like. One may additionally include suitable preservatives, stabilizers, antioxidants, antimicrobials, and buffering agents, for example, BHA, BHT, citric acid, ascorbic acid, tetracycline, and the like. Cream or ointment bases useful in formulation include lanolin, Silvadene® (Marion), Aquaphor®

(Duke Laboratories), and the like. Other topical formulations include aerosols, bandages, and other wound dressings.

Alternatively, one may incorporate or encapsulate the TIAM2 polypeptides, polynucleotides, or antibodies in a suitable polymer matrix or membrane, thus providing a sustained-release delivery device suitable for implantation near the site to be treated locally. Other devices include indwelling catheters and devices such as the Alzet® minipump. Ophthalmic preparations may be formulated using commercially available vehicles such as Sorbi-care® (Allergan), Neodecadron® (Merck, Sharp & Dohme), Lacrilube®, and the like, or may employ topical preparations such as that described in US 5,124,155, incorporated herein by reference. Further, one may provide a TIAM2 polypeptide, polynucleotide, or antibody in solid form, especially as a lyophilized powder. Lyophilized formulations typically contain stabilizing and bulking agents, for example human serum albumin, sucrose, mannitol, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in *Remington's Pharmaceutical Sciences* (Mack Pub. Co.).

The amount of TIAM2 polypeptide, polynucleotide, or antibody required to treat any particular disorder will of course vary depending upon the nature and severity of the disorder, the age and condition of the subject, and other factors readily determined by one of ordinary skill in the art. The appropriate dosage may be determined by one of ordinary skill in the art.

It may be useful to administer the nucleic acid molecules described above, *i.e.*, the ribozyme or antisense molecules, in a gene therapy method. Accordingly, the vectors and techniques described below will be useful. The following expression systems describe vectors, promoters and regulatory elements that are useful for gene therapy applications for the delivery of the above polynucleotides.

Vectors and expression systems useful for the present invention include viral and non-viral systems. Examples of viral delivery systems include retroviruses, adenoviruses, adeno-associated viruses (AAV), sindbis and herpes viruses. In one aspect of the present invention, the viral vector is capable of integrating the above nucleic acid sequence into the host cell genome for long term expression.

One preferred retrovirus is a murine leukemia virus. However, it may be preferred to avoid integration into the host cell genome. Non-viral vectors include naked DNA and DNA formulated with cationic lipids or liposomes. A non-viral system that can be used is the T7/T7 system.

Additionally, it will be useful to produce TIAM2 protein from the presently disclosed nucleic acid sequence to be used in an assay to test for inhibitors or for the preparation of monoclonal antibodies, for example. TIAM2 can be expressed by a baculovirus that has been transformed with a native or modified TIAM2 nucleic acid sequence. The TIAM2 nucleic acid sequence useful in the present invention encodes a protein having an amino acid sequence that is substantially identical to the amino acid sequence of native TIAM2.

Preferably, the TIAM2 nucleic acid or protein sequence will be homologous to the partial sequences listed below. Preferably, the above sequence will be greater than 80% homologous to SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof, more preferably it will be greater than 85% homologous, more preferably greater than 90% homologous, even more preferably greater than 91%, 92%, 93% , 94% or 95% homologous. Most preferably, it will be greater than 96%, 97%, 98% or 99% homologous. Substantial identity means the sequences are identical or differ by one or more alterations (deletion, additions, substitutions) that do not adversely affect the activity of the protein. It is preferable that the protein sequences are homologous in the same percentages noted above. Percent identity or homology can be calculated using methods and algorithms known in the art. A suitable but non-limiting method is the Smith-Waterman homology search algorithm as implemented in MPSRCH program (Oxford Biomolecular) using an affine gap search with the following search parameters: gap open penalty of 12, and gap extension penalty of 1.

The precise chemical structure of the TIAM2 sequence can depend on a number of factors. As ionizable amino and carboxyl groups are present in the molecule, a particular protein may be obtained as a acidic or basic salt, or in neutral form. All such preparations which retain their activity when placed in suitable environmental

conditions are included in the definition of proteins herein. Further, the primary amino acid sequence of the protein may be augmented by derivitization using sugar moieties (glycosylation) or by other supplementary molecules such as lipids, phosphate, acetyl groups and the like. Certain aspects of such augmentation are accomplished through post-translational processing systems of the producing host; other such modifications may be introduced *in vitro*. In any event, such modifications are included in the definition of protein herein so long as the activity of the protein is not destroyed. It is expected that such modifications may quantitatively or qualitatively affect the activity, either by enhancing or diminishing the activity of the protein, in the various assays. Further, individual amino acid residues in the chain may be modified by oxidation, reduction, or derivatization, and the protein may be cleaved to obtain fragments which retain activity. Such alterations which do not destroy activity do not remove the protein sequence from the definition of TIAM2 herein.

Modifications to the primary structure itself, by deletion, addition, or alteration of the amino acids incorporated into the sequence during translation, can be made without destroying the activity of the protein. For example, site specific mutagenesis can enable specific changes in the DNA structure to effect a change in the polypeptide structure. See Mark *et al.* U. S. Pat. No. 4,959,314, and Sambrook, *et al.*, *supra*, Volume 2, chapter 15 which is hereby incorporated by reference in its entirety.

The TIAM2 proteins include mutants, fragments, fusions, and the protein encoded by the sequence listed in SEQ ID NO:1, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof. The native TIAM2 proteins are those that occur in nature. The amino acid sequence of native polypeptides will comprise a sequence that varies slightly; typically, less than by 10-20 amino acids encoded by SEQ ID NO:1, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, or SEQ ID NO:10.

The term "polynucleotide" or "nucleic acid sequence" as used herein refers to a polymer of nucleotides of any length, preferably deoxyribonucleotides, and is used interchangeably herein with the terms "oligonucleotide" and "oligomer." The term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA, as well as antisense polynucleotides. It also includes known

types of modifications, for example, the presence of labels which are known in the art, methylation, end "caps," substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, replacement with certain types of uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) or charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.), introduction of pendant moieties, such as, for example, proteins (including nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (*e.g.*, acridine, psoralen, etc.), chelators (*e.g.*, metals, radioactive species, boron, oxidative moieties, etc.), alkylators (*e.g.*, alpha anomeric nucleic acids, etc.).

A polynucleotide sequence encoding a native TIAM2 protein can be easily modified to encode other classes of TIAM2 proteins. It will be recognized in the art that some amino acid sequence of the TIAM2 polypeptide can be varied without significant effect on the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there are critical areas on the protein which determine activity. In general, it is possible to replace residues that form the tertiary structure, provided that residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the protein. The replacement of amino acids can also change the selectivity of binding to cell surface receptors. Ostade et al., *Nature* *361*:266-268 (1993) describes certain mutations resulting in selective binding of TNF-alpha to only one of the two known types of TNF receptors. Thus, the polypeptides of the present invention may include one or more amino acid substitutions, deletions or additions, either from natural mutations or human manipulation.

Thus, the invention further includes variations of the TIAM2 polypeptide which show substantial TIAM2 polypeptide activity or which include regions of TIAM2 protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions. As indicated above, guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J.U., et al., "Deciphering the Message in Protein Sequences:

Tolerance to Amino Acid Substitutions," *Science 247*:1306-1310 (1990).

Of particular interest are substitutions of charged amino acids with another charged amino acid and with neutral or negatively charged amino acids. The latter results in proteins with reduced positive charge to improve the characteristics of the TIAM2 proteins. The prevention of aggregation is highly desirable. Aggregation of proteins not only results in a loss of activity but can also be problematic when preparing pharmaceutical formulations, because they can be immunogenic. (Pinckard et al., *Clin. Exp. Immunol. 2*:331-340 (1967); Robbins et al., *Diabetes 36*:838-845 (1987); Cleland et al. *Crit. Rev. Therapeutic Drug Carrier Systems 10*:307-377 (1993)).

Amino acids in the polypeptides of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, Science 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or in vitro, or in vitro proliferative activity. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith et al., *J. Mol. Biol. 224*:899-904 (1992) and de Vos et al. *Science 255*:306-312 (1992)).

For example, mutants can be constructed by making conservative amino acid substitutions. The following are examples of conservative substitutions: Gly √ Ala; Val √ Ile √ Leu; Asp √ Glu; Lys √ Arg; Asn √ Gln; and Phe √ Trp √ Tyr. A subset of mutants, called muteins, is a group of polypeptides with the non-disulfide bond participating cysteines substituted with a neutral amino acid, generally, with serines. These mutants may be stable over a broader temperature range than native TIAM2 proteins. As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein. Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of substitutions for any given TIAM2 polypeptide will not be more than 50, 40,

30, 25, 20, 15, 10, 5 or 3. The coding sequence of mutants can be constructed by *in vitro* mutagenesis of the native TIAM2 polypeptide coding sequences.

As used herein, the term "protein" or "polypeptide" refers to a polymer of amino acids and does not refer to a specific length of the product; thus, peptides, oligopeptides, polypeptides, proteins, and polyproteins, as well as fragments of these, are included within this definition. This term also does not refer to, or exclude, post expression modifications of the protein, for example, glycosylations, acetylations, phosphorylations and the like. Included within the definition are, for example, proteins containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), proteins with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

A polypeptide or protein or amino acid sequence "derived from" or "coded by" or "encoded by" a designated nucleic acid sequence refers to a polypeptide having an amino acid sequence identical to that of a polypeptide encoded in the sequence, or a portion thereof wherein the portion consists of at least 3-5 consecutive amino acids, and more preferably at least 8-10 amino acids, and even more preferably at least 11-15 amino acids, or which is immunologically identifiable with a polypeptide encoded in the sequence. This terminology also includes a polypeptide expressed from a designated nucleic acid sequence.

"Purified" and "isolated" mean, when referring to a polypeptide or nucleotide sequence, that the indicated molecule is present in the substantial absence of other biological macromolecules of the same type. The term "purified" as used herein preferably means at least 75% by weight, more preferably at least 85% by weight, more preferably still at least 95% by weight, and most preferably at least 98% 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 1000, can be present).

Also included within the scope of the invention are fragments of polynucleotides that encode the TIAM2 polypeptides disclosed herein. Such fragments may encode antigenic regions of the proteins or biologically functional regions of the

proteins as described herein. Such fragments also encode polypeptide components of the fusion proteins disclosed herein. Fragments are also suitable for use as probes for identifying polynucleotides capable of hybridizing with SEQ ID NO:8, 9, and/or 10. Fragments can range in length from about 10 basepairs to the full length polynucleotide, which is 3344 basepairs for SEQ ID NO:8, 4590 basepairs for SEQ ID NO:9, and 4514 basepairs for SEQ ID NO:10. Preferred fragments are 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, or 3300 basepairs for SEQ ID NO: 8, 9, or 10; 3400, 3500, 3600, 3700, 3800, 4000, 4100, 4200, 4300, 4400, 4500, or 4510 basepairs for SEQ ID NO:9 or 10; and 4520, 4550, 4575, or 4580 basepairs for SEQ ID NO:10. A particularly preferred fragment comprises nucleotides 1-104 of SEQ ID NO:8.

Polypeptide fragments are amino and/or carboxyl terminal amino acid deletions of mutant or native TIAM2 proteins. The number of amino acids that are truncated is not critical as long as the polypeptide fragment exhibits the desired sequence homology, immunological or biological activity. Preferred fragments are 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600 or 625 amino acids of SEQ ID NO:7, 11, 12 or 13; 650, 675, 700, 750, 800, 825, 900, 950, 1000, 1025, or 1050 amino acids of SEQ ID NO:12 or 13; and 1060, 1065, 1070 or 1075 amino acids of SEQ ID NO:13. Polypeptide fragments of immunological significance comprise, for example, at least one epitope shared by a native TIAM2 protein. Such TIAM2 proteins may be only 5-15 amino acids in length. Examples of amino acid sequence of fragments comprise amino acid number 1-8 (aa1 to aa8) of SEQ ID NO:7, 11, 12, or 13; aa2 to aa9 of SEQ ID NO:7, 11, 12, or 13; aa3 to aa10 of SEQ ID NO:7, 11, 12, or 13; aa4 to aa11 of SEQ ID NO:7, 11, 12, or 13; aa5 to aa12 of SEQ ID NO:7, 11, 12, or 13; aa6 to aa13 of SEQ ID NO:7, 11, 12, or 13; aa7 to aa14 of SEQ ID NO:7, 11, 12, or 13; aa8 to aa15 of SEQ ID NO:7, 11, 12, or 13; aa9 to aa16 of SEQ ID NO:7, 11, 12, or 13; aa10 to aa17 of SEQ ID NO:7, 11, 12, or 13; aa11 to aa18 of SEQ ID NO:7, 11, 12, or 13; aa12 to aa19 of

SEQ ID NO:7, 11, 12, or 13; aa13 to aa20 of SEQ ID NO:7, 11, 12, or 13: aa14 to aa21 of SEQ ID NO:7, 11, 12, or 13; aa15 to aa22 of SEQ ID NO:7, 11, 12, or 13; aa16 to aa23 of SEQ ID NO:7, 11, 12, or 13; aa17 to aa24 of SEQ ID NO:7, 11, 12, or 13; aa18 to aa25 of SEQ ID NO:7, 11, 12, or 13; aa19 to aa26 of SEQ ID NO:7, 11, 12, or 13; aa20 to aa27 of SEQ ID NO:7, 11, 12, or 13; aa21 to aa28 of SEQ ID NO:7, 11, 12, or 13; aa22 to aa29 of SEQ ID NO:7, 11, 12, or 13; aa23 to aa30 of SEQ ID NO:7, 11, 12, or 13; aa24 to aa31 of SEQ ID NO:7, 11, 12, or 13; aa25 to aa32 of SEQ ID NO:7, 11, 12, or 13; aa26 to aa33 of SEQ ID NO:7, 11, 12, or 13; aa27 to aa34 of SEQ ID NO:7, 11, 12, or 13; aa28 to aa35 of SEQ ID NO:7, 11, 12, or 13; aa29 to aa36 of SEQ ID NO:7, 11, 12, or 13; aa30 to aa37 of SEQ ID NO:7, 11, 12, or 13; aa31 to aa38 of SEQ ID NO:7, 11, 12, or 13; aa32 to aa39 of SEQ ID NO:7, 11, 12, or 13; aa33 to aa40 of SEQ ID NO:7, 11, 12, or 13; aa34 to aa41 of SEQ ID NO:7, 11, 12, or 13; aa35 to aa42 of SEQ ID NO:7, 11, 12, or 13; aa36 to aa43 of SEQ ID NO:7, 11, 12, or 13; aa37 to aa44 of SEQ ID NO:7, 11, 12, or 13, aa38 to aa45 of SEQ ID NO:7, 11, 12, or 13; aa39 to aa46 of SEQ ID NO:7, 11, 12, or 13; aa40 to aa47 of SEQ ID NO:7, 11, 12, or 13; aa41 to aa48 of SEQ ID NO:7, 11, 12, or 13; aa42 to aa49 of SEQ ID NO:7, 11, 12, or 13; aa43 to aa50 of SEQ ID NO:7, 11, 12, or 13; aa44 to aa51 of SEQ ID NO:7, 11, 12, or 13; aa45 to aa52 of SEQ ID NO:7, 11, 12, or 13; aa46 to aa53 of SEQ ID NO:7, 11, 12, or 13; aa47 to aa54 of SEQ ID NO:7, 11, 12, or 13; aa48 to aa55 of SEQ ID NO:7, 11, 12, or 13; aa49 to aa56 of SEQ ID NO:7, 11, 12, or 13; aa50 to aa57 of SEQ ID NO:7, 11, 12, or 13; aa51 to aa58 of SEQ ID NO:7, 11, 12, or 13; aa52 to aa59 of SEQ ID NO:7, 11, 12, or 13; aa53 to aa60 of SEQ ID NO:7, 11, 12, or 13; aa54 to aa61 of SEQ ID NO:7, 11, 12, or 13; aa55 to aa62 of SEQ ID NO:7, 11, 12, or 13; aa56 to aa63 of SEQ ID NO:7, 11, 12, or 13; aa57 to aa64 of SEQ ID NO:7, 11, 12, or 13; aa58 to aa65 of SEQ ID NO:7, 11, 12, or 13; aa59 to aa66 of SEQ ID NO:7, 11, 12, or 13; aa60 to aa67 of SEQ ID NO:7, 11, 12, or 13; aa61 to aa68 of SEQ ID NO:7, 11, 12, or 13; aa62 to aa69 of SEQ ID NO:7, 11, 12, or 13; aa63 to aa70 of SEQ ID NO:7, 11, 12, or 13; aa64 to aa71 of SEQ ID NO:7, 11, 12, or 13; aa65 to aa72 of SEQ ID NO:7, 11, 12, or 13; aa66 to aa73 of SEQ ID NO:7, 11, 12, or 13; aa67 to aa74 of SEQ ID NO:7, 11, 12, or 13; aa68 to aa75 of SEQ ID NO:7, 11, 12, or 13; aa69 to aa76 of SEQ ID NO:7,

11, 12, or 13; aa70 to aa77 of SEQ ID NO:7, 11, 12, or 13; aa71 to aa78 of SEQ ID NO:7, 11, 12, or 13; aa72 to aa79 of SEQ ID NO:7, 11, 12, or 13; aa73 to aa80 of SEQ ID NO:7, 11, 12, or 13; aa74 to aa81 of SEQ ID NO:7, 11, 12, or 13; aa75 to aa82 of SEQ ID NO:7, 11, 12, or 13; aa76 to aa83 of SEQ ID NO:7, 11, 12, or 13; aa77 to aa84 of SEQ ID NO:7, 11, 12, or 13; aa78 to aa84 of SEQ ID NO:7, 11, 12, or 13; aa79 to aa86 of SEQ ID NO:7, 11, 12, or 13; aa80 to aa87 of SEQ ID NO:7, 11, 12, or 13; aa81 to aa88 of SEQ ID NO:7, 11, 12, or 13; aa82 to aa89 of SEQ ID NO:7, 11, 12, or 13; aa83 to aa90 of SEQ ID NO:7, 11, 12, or 13; aa84 to aa91 of SEQ ID NO:7, 11, 12, or 13; aa85 to aa92 of SEQ ID NO:7, 11, 12, or 13; aa86 to aa93 of SEQ ID NO:7, 11, 12, or 13; aa87 to aa94 of SEQ ID NO:7, 11, 12, or 13; aa88 to aa95 of SEQ ID NO:7, 11, 12, or 13; aa89 to aa96 of SEQ ID NO:7, 11, 12, or 13; aa90 to aa97 of SEQ ID NO:7, 11, 12, or 13; aa91 to aa98 of SEQ ID NO:7, 11, 12, or 13; aa92 to aa99 of SEQ ID NO:7, 11, 12, or 13; aa93 to aa100 of SEQ ID NO:7, 11, 12, or 13; aa94 to aa101 of SEQ ID NO:7, 11, 12, or 13; aa95 to aa102 of SEQ ID NO:7, 11, 12, or 13; aa96 to aa103 of SEQ ID NO:7, 11, 12, or 13; aa97 to aa104 of SEQ ID NO:7, 11, 12, or 13; aa98 to aa105 of SEQ ID NO:7, 11, 12, or 13; aa99 to aa106 of SEQ ID NO:7, 11, 12, or 13; aa100 to aa107 of SEQ ID NO:7, 11, 12, or 13 aa101 to aa108 of SEQ ID NO:7, 11, 12, or 13 aa102 to aa109 of SEQ ID NO:7, 11, 12, or 13; aa103 to aa110 of SEQ ID NO:7, 11, 12, or 13; aa104 to aa111 of SEQ ID NO:7, 11, 12, or 13; aa105 to aa112 of SEQ ID NO:7, 11, 12, or 13; aa106 to aa113 of SEQ ID NO:7, 11, 12, or 13; aa107 to aa114 of SEQ ID NO:7, 11, 12, or 13; aa108 to aa115 of SEQ ID NO:7, 11, 12, or 13; aa109 to aa116 of SEQ ID NO:7, 11, 12, or 13; aa110 to aa117 of SEQ ID NO:7, 11, 12, or 13; aa111 to aa118 of SEQ ID NO:7, 11, 12, or 13; aa112 to aa119 of SEQ ID NO:7, 11, 12, or 13; aa113 to aa120 of SEQ ID NO:7, 11, 12, or 13; aa114 to aa121 of SEQ ID NO:7, 11, 12, or 13; aa115 to aa122 of SEQ ID NO:7, 11, 12, or 13; aa116 to aa123 of SEQ ID NO:7, 11, 12, or 13; aa117 to aa124 of SEQ ID NO:7, 11, 12, or 13; aa118 to aa125 of SEQ ID NO:7, 11, 12, or 13; aa119 to aa126 of SEQ ID NO:7, 11, 12, or 13; aa120 to aa127 of SEQ ID NO:7, 11, 12, or 13; aa121 to aa128 of SEQ ID NO:7, 11, 12, or 13; aa122 to aa129 of SEQ ID NO:7, 11, 12, or 13; aa123 to aa130 of SEQ ID NO:7, 11, 12, or 13; aa124 to aa131 of SEQ ID NO:7, 11, 12, or 13; aa125 to

aa132 of SEQ ID NO:7, 11, 12, or 13; aa126 to aa133 of SEQ ID NO:7, 11, 12, or 13; aa127 to aa134 of SEQ ID NO:7, 11, 12, or 13; aa128 to aa135 of SEQ ID NO:7, 11, 12, or 13; aa129 to aa136 of SEQ ID NO:7, 11, 12, or 13; aa130 to aa137 of SEQ ID NO:7, 11, 12, or 13; aa131 to aa138 of SEQ ID NO:7, 11, 12, or 13; aa132 to aa139 of SEQ ID NO:7, 11, 12, or 13; aa133 to aa140 of SEQ ID NO:7, 11, 12, or 13; aa134 to aa141 of SEQ ID NO:7, 11, 12, or 13; aa135 to aa142 of SEQ ID NO:7, 11, 12, or 13; aa136 to aa143 of SEQ ID NO:7, 11, 12, or 13; aa137 to aa144 of SEQ ID NO:7, 11, 12, or 13; aa138 to aa145 of SEQ ID NO:7, 11, 12, or 13; aa139 to aa146 of SEQ ID NO:7, 11, 12, or 13; aa140 to aa147 of SEQ ID NO:7, 11, 12, or 13; aa141 to aa148 of SEQ ID NO:7, 11, 12, or 13; aa142 to aa149 of SEQ ID NO:7, 11, 12, or 13; aa143 to aa150 of SEQ ID NO:7, 11, 12, or 13; aa144 to aa151 of SEQ ID NO:7, 11, 12, or 13; aa145 to aa152 of SEQ ID NO:7, 11, 12, or 13; aa146 to aa153 of SEQ ID NO:7, 11, 12, or 13; aa147 to aa154 of SEQ ID NO:7, 11, 12, or 13; aa148 to aa155 of SEQ ID NO:7, 11, 12, or 13; aa149 to aa156 of SEQ ID NO:7, 11, 12, or 13; aa150 to aa157 of SEQ ID NO:7, 11, 12, or 13; aa151 to aa158 of SEQ ID NO:7, 11, 12, or 13; aa152 to aa159 of SEQ ID NO:7, 11, 12, or 13; aa153 to aa160 of SEQ ID NO:7, 11, 12, or 13; aa154 to aa161 of SEQ ID NO:7, 11, 12, or 13; aa155 to aa162 of SEQ ID NO:7, 11, 12, or 13; aa156 to aa163 of SEQ ID NO:7, 11, 12, or 13; aa157 to aa164 of SEQ ID NO:7, 11, 12, or 13; aa158 to aa165 of SEQ ID NO:7, 11, 12, or 13; aa159 to aa166 of SEQ ID NO:7, 11, 12, or 13; aa160 to aa167 of SEQ ID NO:7, 11, 12, or 13; aa161 to aa168 of SEQ ID NO:7, 11, 12, or 13; aa162 to aa169 of SEQ ID NO:7, 11, 12, or 13; aa163 to aa170 of SEQ ID NO:7, 11, 12, or 13; aa164 to aa171 of SEQ ID NO:7, 11, 12, or 13; aa165 to aa172 of SEQ ID NO:7, 11, 12, or 13; aa166 to aa173 of SEQ ID NO:7, 11, 12, or 13; aa167 to aa174 of SEQ ID NO:7, 11, 12, or 13; aa168 to aa175 of SEQ ID NO:7, 11, 12, or 13; aa169 to aa176 of SEQ ID NO:7, 11, 12, or 13; aa170 to aa177 of SEQ ID NO:7, 11, 12, or 13; aa171 to aa178 of SEQ ID NO:7, 11, 12, or 13; aa172 to aa179 of SEQ ID NO:7, 11, 12, or 13; aa173 to aa180 of SEQ ID NO:7, 11, 12, or 13; aa174 to aa181 of SEQ ID NO:7, 11, 12, or 13; aa175 to aa182 of SEQ ID NO:7, 11, 12, or 13; aa176 to aa183 of SEQ ID NO:7, 11, 12, or 13; aa177 to aa184 of SEQ ID NO:7, 11, 12, or 13; aa178 to aa185 of SEQ ID NO:7, 11, 12, or 13; aa179 to

aa186 of SEQ ID NO:7, 11, 12, or 13; aa180 to aa187 of SEQ ID NO:7, 11, 12, or 13; aa181 to aa188 of SEQ ID NO:7, 11, 12, or 13; aa182 to aa189 of SEQ ID NO:7, 11, 12, or 13; aa183 to aa190 of SEQ ID NO:7, 11, 12, or 13; aa184 to aa191 of SEQ ID NO:7, 11, 12, or 13; aa185 to aa192 of SEQ ID NO:7, 11, 12, or 13; aa186 to aa193 of SEQ ID NO:7, 11, 12, or 13; aa187 to aa194 of SEQ ID NO:7, 11, 12, or 13; aa188 to aa195 of SEQ ID NO:7, 11, 12, or 13; aa189 to aa196 of SEQ ID NO:7, 11, 12, or 13; aa190 to aa197 of SEQ ID NO:7, 11, 12, or 13; aa191 to aa198 of SEQ ID NO:7, 11, 12, or 13; aa192 to aa199 of SEQ ID NO:7, 11, 12, or 13; aa193 to aa200 of SEQ ID NO:7, 11, 12, or 13; aa194 to aa201 of SEQ ID NO:7, 11, 12, or 13; aa195 to aa202 of SEQ ID NO:7, 11, 12, or 13; aa196 to aa203 of SEQ ID NO:7, 11, 12, or 13; aa197 to aa204 of SEQ ID NO:7, 11, 12, or 13; aa198 to aa205 of SEQ ID NO:7, 11, 12, or 13; aa199 to aa206 of SEQ ID NO:7, 11, 12, or 13; aa200 to aa207 of SEQ ID NO:7, 11, 12, or 13; aa201 to aa208 of SEQ ID NO:7, 11, 12, or 13; aa202 to aa209 of SEQ ID NO:7, 11, 12, or 13; aa203 to aa210 of SEQ ID NO:7, 11, 12, or 13; aa204 to aa211 of SEQ ID NO:7, 11, 12, or 13; aa205 to aa212 of SEQ ID NO:7, 11, 12, or 13; aa206 to aa213 of SEQ ID NO:7, 11, 12, or 13; aa207 to aa214 of SEQ ID NO:7, 11, 12, or 13; aa208 to aa215 of SEQ ID NO:7, 11, 12, or 13; aa209 to aa216 of SEQ ID NO:7, 11, 12, or 13; aa210 to aa217 of SEQ ID NO:7, 11, 12, or 13; aa211 to aa218 of SEQ ID NO:7, 11, 12, or 13; aa212 to aa219 of SEQ ID NO:7, 11, 12, or 13; aa213 to aa220 of SEQ ID NO:7, 11, 12, or 13; aa214 to aa221 of SEQ ID NO:7, 11, 12, or 13; aa215 to aa222 of SEQ ID NO:7, 11, 12, or 13; aa216 to aa223 of SEQ ID NO:7, 11, 12, or 13; aa217 to aa224 of SEQ ID NO:7, 11, 12, or 13; aa218 to aa225 of SEQ ID NO:7, 11, 12, or 13; aa219 to aa226 of SEQ ID NO:7, 11, 12, or 13; aa220 to aa227 of SEQ ID NO:7, 11, 12, or 13; aa221 to aa228 of SEQ ID NO:7, 11, 12, or 13; aa222 to aa229 of SEQ ID NO:7, 11, 12, or 13; aa223 to aa230 of SEQ ID NO:7, 11, 12, or 13; aa224 to aa231 of SEQ ID NO:7, 11, 12, or 13; aa225 to aa232 of SEQ ID NO:7, 11, 12, or 13; and aa226 to aa233; of SEQ ID NO:7, 11, 12, or 13 aa227 to aa234 of SEQ ID NO:7, 11, 12, or 13; aa228 to aa235 of SEQ ID NO:7, 11, 12, or 13; aa229 to aa236 of SEQ ID NO:7, 11, 12, or 13; aa230 to aa237 of SEQ ID NO:7, 11, 12, or 13; aa231 to aa238 of SEQ ID NO:7, 11, 12, or 13; aa222 to aa239 of SEQ ID NO:7, 11, 12, or 13; aa233 to

aa240 of SEQ ID NO:7, 11, 12, or 13; aa234 to aa41 of SEQ ID NO:7, 11, 12, or 13; aa235 to aa242 of SEQ ID NO:7, 11, 12, or 13; aa236 to aa243 of SEQ ID NO:7, 11, 12, or 13; aa237 to aa244 of SEQ ID NO:7, 11, 12, or 13; aa238 to aa245 of SEQ ID NO:7, 11, 12, or 13; aa239 to aa246 of SEQ ID NO:7, 11, 12, or 13; aa240 to aa247 of SEQ ID NO:7, 11, 12, or 13; aa241 to aa248 of SEQ ID NO:7, 11, 12, or 13; aa242 to aa249 of SEQ ID NO:7, 11, 12, or 13; aa243 to aa250 of SEQ ID NO:7, 11, 12, or 13; aa244 to aa251 of SEQ ID NO:7, 11, 12, or 13; aa245 to aa252 of SEQ ID NO:7, 11, 12, or 13; aa246 to aa253 of SEQ ID NO:7, 11, 12, or 13; aa247 to aa254 of SEQ ID NO:7, 11, 12, or 13; aa248 to aa255 of SEQ ID NO:7, 11, 12, or 13; aa249 to aa256 of SEQ ID NO:7, 11, 12, or 13; aa250 to aa257 of SEQ ID NO:7, 11, 12, or 13; aa251 to aa258 of SEQ ID NO:7, 11, 12, or 13; aa252 to aa259 of SEQ ID NO:7, 11, 12, or 13; aa253 to aa260 of SEQ ID NO:7, 11, 12, or 13; aa254 to aa261 of SEQ ID NO:7, 11, 12, or 13; aa255 to aa262 of SEQ ID NO:7, 11, 12, or 13; aa256 to aa263 of SEQ ID NO:7, 11, 12, or 13; aa257 to aa264 of SEQ ID NO:7, 11, 12, or 13; aa258 to aa265 of SEQ ID NO:7, 11, 12, or 13; aa259 to aa266 of SEQ ID NO:7, 11, 12, or 13; aa260 to aa267 of SEQ ID NO:7, 11, 12, or 13; aa261 to aa268 of SEQ ID NO:7, 11, 12, or 13; aa262 to aa269 of SEQ ID NO:7, 11, 12, or 13; aa263 to aa270 of SEQ ID NO:7, 11, 12, or 13; aa264 to aa271 of SEQ ID NO:7, 11, 12, or 13; aa265 to aa272 of SEQ ID NO:7, 11, 12, or 13; aa266 to aa273 of SEQ ID NO:7, 11, 12, or 13; aa267 to aa274 of SEQ ID NO:7, 11, 12, or 13; aa268 to aa275 of SEQ ID NO:7, 11, 12, or 13; aa269 to aa276 of SEQ ID NO:7, 11, 12, or 13; aa270 to aa277 of SEQ ID NO:7, 11, 12, or 13; aa271 to aa278 of SEQ ID NO:7, 11, 12, or 13; aa272 to aa279 of SEQ ID NO:7, 11, 12, or 13; aa273 to aa280 of SEQ ID NO:7, 11, 12, or 13; aa274 to aa281 of SEQ ID NO:7, 11, 12, or 13; aa275 to aa282 of SEQ ID NO:7, 11, 12, or 13; aa276 to aa283 of SEQ ID NO:7, 11, 12, or 13; aa277 to aa284 of SEQ ID NO:7, 11, 12, or 13; aa278 to aa284 of SEQ ID NO:7, 11, 12, or 13; aa279 to aa286 of SEQ ID NO:7, 11, 12, or 13; aa280 to aa287 of SEQ ID NO:7, 11, 12, or 13; aa281 to aa288 of SEQ ID NO:7, 11, 12, or 13; aa282 to aa289 of SEQ ID NO:7, 11, 12, or 13; aa283 to aa290 of SEQ ID NO:7, 11, 12, or 13; aa284 to aa291 of SEQ ID NO:7, 11, 12, or 13; aa285 to aa292 of SEQ ID NO:7, 11, 12, or 13; aa286 to aa293 of SEQ ID NO:7, 11, 12, or 13; aa287 to

aa294 of SEQ ID NO:7, 11, 12, or 13; aa288 to aa295 of SEQ ID NO:7, 11, 12, or 13; aa289 to aa296 of SEQ ID NO:7, 11, 12, or 13; aa290 to aa297 of SEQ ID NO:7, 11, 12, or 13; aa291 to aa298 of SEQ ID NO:7, 11, 12, or 13; aa292 to aa299 of SEQ ID NO:7, 11, 12, or 13; aa293 to aa300 of SEQ ID NO:7, 11, 12, or 13; aa294 to aa101 of SEQ ID NO:7, 11, 12, or 13; aa295 to aa102 of SEQ ID NO:7, 11, 12, or 13; aa296 to aa103 of SEQ ID NO:7, 11, 12, or 13; aa297 to aa104 of SEQ ID NO:7, 11, 12, or 13; aa298 to aa105 of SEQ ID NO:7, 11, 12, or 13; aa299 to aa106 of SEQ ID NO:7, 11, 12, or 13; aa300 to aa307 of SEQ ID NO:7, 11, 12, or 13 aa301 to aa308 of SEQ ID NO:7, 11, 12, or 13 aa302 to aa309 of SEQ ID NO:7, 11, 12, or 13; aa303 to aa310 of SEQ ID NO:7, 11, 12, or 13; aa304 to aa311 of SEQ ID N; O:7; aa305 to aa312 of SEQ ID NO:7, 11, 12, or 13; aa306 to aa313 of SEQ ID NO:7, 11, 12, or 13; aa307 to aa314 of SEQ ID NO:7, 11, 12, or 13; aa308 to aa315 of SEQ ID NO:7, 11, 12, or 13; aa309 to aa316 of SEQ ID NO:7, 11, 12, or 13; aa310 to aa317 of SEQ ID NO:7, 11, 12, or 13aa311 to aa318 of SEQ ID NO:7, 11, 12, or 13; aa312 to aa319 of SEQ ID NO:7, 11, 12, or 13; aa313 to aa320 of SEQ ID NO:7, 11, 12, or 13; aa314 to aa321 of SEQ ID NO:7, 11, 12, or 13; aa315 to aa322 of SEQ ID NO:7, 11, 12, or 13; aa316 to aa323 of SEQ ID NO:7, 11, 12, or 13; aa317 to aa324 of SEQ ID NO:7, 11, 12, or 13; aa318 to aa325 of SEQ ID NO:7, 11, 12, or 13; aa319 to aa326 of SEQ ID NO:7, 11, 12, or 13; aa320 to aa327 of SEQ ID NO:7, 11, 12, or 13; aa321 to aa328 of SEQ ID NO:7, 11, 12, or 13; aa322 to aa329 of SEQ ID NO:7, 11, 12, or 13; aa323 to aa330 of SEQ ID NO:7, 11, 12, or 13; aa324 to aa331 of SEQ ID NO:7, 11, 12, or 13; aa325 to aa332 of SEQ ID NO:7, 11, 12, or 13; aa326 to aa333 of SEQ ID NO:7, 11, 12, or 13; aa327 to aa334 of SEQ ID NO:7, 11, 12, or 13; aa328 to aa335 of SEQ ID NO:7, 11, 12, or 13; aa329 to aa336 of SEQ ID NO:7, 11, 12, or 13; aa330 to aa337 of SEQ ID NO:7, 11, 12, or 13; aa331 to aa338 of SEQ ID NO:7, 11, 12, or 13; aa332 to aa339 of SEQ ID NO:7, 11, 12, or 13; aa333 to aa340 of SEQ ID NO:7, 11, 12, or 13; aa334 to aa341 of SEQ ID NO:7, 11, 12, or 13; aa335 to aa342 of SEQ ID NO:7, 11, 12, or 13; aa336 to aa343 of SEQ ID NO:7, 11, 12, or 13; aa337 to aa344 of SEQ ID NO:7, 11, 12, or 13; aa338 to aa345 of SEQ ID NO:7, 11, 12, or 13; aa339 to aa346 of SEQ ID NO:7, 11, 12, or 13; aa340 to aa347 of SEQ ID NO:7, 11, 12, or 13; aa341 to aa348 of SEQ ID

NO:7, 11, 12, or 13; aa342 to aa349 of SEQ ID NO:7, 11, 12, or 13; aa343 to aa350 of SEQ ID NO:7, 11, 12, or 13; aa344 to aa351 of SEQ ID NO:7, 11, 12, or 13; aa345 to aa352 of SEQ ID NO:7, 11, 12, or 13; aa346 to aa353 of SEQ ID NO:7, 11, 12, or 13; aa347 to aa354 of SEQ ID NO:7, 11, 12, or 13; aa348 to aa355 of SEQ ID NO:7, 11, 12, or 13; aa349 to aa356 of SEQ ID NO:7, 11, 12, or 13; aa350 to aa357 of SEQ ID NO:7, 11, 12, or 13; aa351 to aa358 of SEQ ID NO:7, 11, 12, or 13; aa352 to aa359 of SEQ ID NO:7, 11, 12, or 13; aa353 to aa360 of SEQ ID NO:7, 11, 12, or 13; aa354 to aa361 of SEQ ID NO:7, 11, 12, or 13; aa355 to aa362 of SEQ ID NO:7, 11, 12, or 13; aa356 to aa363 of SEQ ID NO:7, 11, 12, or 13; aa357 to aa364 of SEQ ID NO:7, 11, 12, or 13; aa358 to aa365 of SEQ ID NO:7, 11, 12, or 13; aa359 to aa366 of SEQ ID NO:7, 11, 12, or 13; aa360 to aa367 of SEQ ID NO:7, 11, 12, or 13; aa361 to aa368 of SEQ ID NO:7, 11, 12, or 13; aa362 to aa369 of SEQ ID NO:7, 11, 12, or 13; aa363 to aa370 of SEQ ID NO:7, 11, 12, or 13; aa364 to aa371 of SEQ ID NO:7, 11, 12, or 13; aa365 to aa372 of SEQ ID NO:7, 11, 12, or 13; aa366 to aa373 of SEQ ID NO:7, 11, 12, or 13; aa367 to aa374 of SEQ ID NO:7, 11, 12, or 13; aa368 to aa375 of SEQ ID NO:7, 11, 12, or 13; aa369 to aa376 of SEQ ID NO:7, 11, 12, or 13; aa370 to aa377 of SEQ ID NO:7, 11, 12, or 13; aa371 to aa378 of SEQ ID NO:7, 11, 12, or 13; aa372 to aa379 of SEQ ID NO:7, 11, 12, or 13; aa373 to aa380 of SEQ ID NO:7, 11, 12, or 13; aa374 to aa381 of SEQ ID NO:7, 11, 12, or 13; aa375 to aa382 of SEQ ID NO:7, 11, 12, or 13; aa376 to aa383 of SEQ ID NO:7, 11, 12, or 13; aa377 to aa384 of SEQ ID NO:7, 11, 12, or 13; aa378 to aa385 of SEQ ID NO:7, 11, 12, or 13; aa379 to aa386 of SEQ ID NO:7, 11, 12, or 13; aa380 to aa387 of SEQ ID NO:7, 11, 12, or 13; aa381 to aa388 of SEQ ID NO:7, 11, 12, or 13; aa382 to aa389 of SEQ ID NO:7, 11, 12, or 13; aa383 to aa390 of SEQ ID NO:7, 11, 12, or 13; aa384 to aa391 of SEQ ID NO:7, 11, 12, or 13; aa385 to aa392 of SEQ ID NO:7, 11, 12, or 13; aa386 to aa393 of SEQ ID NO:7, 11, 12, or 13; aa387 to aa394 of SEQ ID NO:7, 11, 12, or 13; aa388 to aa395 of SEQ ID NO:7, 11, 12, or 13; aa389 to aa396 of SEQ ID NO:7, 11, 12, or 13; aa390 to aa397 of SEQ ID NO:7, 11, 12, or 13; aa391 to aa398 of SEQ ID NO:7, 11, 12, or 13; aa392 to aa399 of SEQ ID NO:7, 11, 12, or 13; aa393 to aa400 of SEQ ID NO:7, 11, 12, or 13; aa394 to aa401 of SEQ ID NO:7, 11, 12, or 13; aa395 to aa402 of SEQ ID

NO:7, 11, 12, or 13; aa396 to aa403 of SEQ ID NO:7, 11, 12, or 13; aa397 to aa404 of SEQ ID NO:7, 11, 12, or 13; aa398 to aa405 of SEQ ID NO:7, 11, 12, or 13; aa399 to aa406 of SEQ ID NO:7, 11, 12, or 13; aa400 to aa407 of SEQ ID NO:7, 11, 12, or 13; aa401 to aa408 of SEQ ID NO:7, 11, 12, or 13; aa402 to aa409 of SEQ ID NO:7, 11, 12, or 13; aa403 to aa410 of SEQ ID NO:7, 11, 12, or 13; aa404 to aa411 of SEQ ID NO:7, 11, 12, or 13; aa405 to aa412 of SEQ ID NO:7, 11, 12, or 13; aa406 to aa413 of SEQ ID NO:7, 11, 12, or 13; aa407 to aa414 of SEQ ID NO:7, 11, 12, or 13; aa408 to aa415 of SEQ ID NO:7, 11, 12, or 13; aa409 to aa416 of SEQ ID NO:7, 11, 12, or 13; aa410 to aa417 of SEQ ID NO:7, 11, 12, or 13; aa411 to aa418 of SEQ ID NO:7, 11, 12, or 13; aa412 to aa419 of SEQ ID NO:7, 11, 12, or 13; aa413 to aa420 of SEQ ID NO:7, 11, 12, or 13; aa414 to aa421 of SEQ ID NO:7, 11, 12, or 13; aa415 to aa422 of SEQ ID NO:7, 11, 12, or 13; aa416 to aa423 of SEQ ID NO:7, 11, 12, or 13; aa417 to aa424 of SEQ ID NO:7, 11, 12, or 13; aa418 to aa425 of SEQ ID NO:7, 11, 12, or 13; aa419 to aa426 of SEQ ID NO:7, 11, 12, or 13; aa420 to aa427 of SEQ ID NO:7, 11, 12, or 13; aa421 to aa428 of SEQ ID NO:7, 11, 12, or 13; aa422 to aa429 of SEQ ID NO:7, 11, 12, or 13, aa423 to aa430 of SEQ ID NO:7, 11, 12, or 13; aa424 to aa431 of SEQ ID NO:7, 11, 12, or 13; aa425 to aa432 of SEQ ID NO:7, 11, 12, or 13; aa426 to aa433 of SEQ ID NO:7, 11, 12, or 13; aa427 to aa434 of SEQ ID NO:7, 11, 12, or 13; aa428 to aa435 of SEQ ID NO:7, 11, 12, or 13; aa429 to aa436 of SEQ ID NO:7, 11, 12, or 13; aa430 to aa437 of SEQ ID NO:7, 11, 12, or 13; aa431 to aa438 of SEQ ID NO:7, 11, 12, or 13; aa432 to aa439 of SEQ ID NO:7, 11, 12, or 13; aa433 to aa440 of SEQ ID NO:7, 11, 12, or 13; aa434 to aa441 of SEQ ID NO:7, 11, 12, or 13; aa435 to aa442 of SEQ ID NO:7, 11, 12, or 13; aa436 to aa443 of SEQ ID NO:7, 11, 12, or 13; aa437 to aa444 of SEQ ID NO:7, 11, 12, or 13; aa438 to aa445 of SEQ ID NO:7, 11, 12, or 13; aa439 to aa446 of SEQ ID NO:7, 11, 12, or 13; aa440 to aa447 of SEQ ID NO:7, 11, 12, or 13; aa441 to aa448 of SEQ ID NO:7, 11, 12, or 13; aa442 to aa449 of SEQ ID NO:7, 11, 12, or 13; aa443 to aa450 of SEQ ID NO:7, 11, 12, or 13; aa444 to aa451 of SEQ ID NO:7, 11, 12, or 13; aa445 to aa452 of SEQ ID NO:7, 11, 12, or 13; aa446 to aa453 of SEQ ID NO:7, 11, 12, or 13; aa447 to aa454 of SEQ ID NO:7, 11, 12, or 13; aa448 to aa455 of SEQ ID NO:7, 11, 12, or 13; aa449 to aa456 of SEQ ID

NO:7, 11, 12, or 13; aa450 to aa457 of SEQ ID NO:7, 11, 12, or 13; aa451 to aa458 of SEQ ID NO:7, 11, 12, or 13; aa452 to aa459 of SEQ ID NO:7, 11, 12, or 13; aa453 to aa460 of SEQ ID NO:7, 11, 12, or 13; aa454 to aa461 of SEQ ID NO:7, 11, 12, or 13; aa455 to aa462 of SEQ ID NO:7, 11, 12, or 13; aa456 to aa463 of SEQ ID NO:7, 11, 12, or 13; aa457 to aa464 of SEQ ID NO:7, 11, 12, or 13; aa458 to aa465 of SEQ ID NO:7, 11, 12, or 13; aa459 to aa466 of SEQ ID NO:7, 11, 12, or 13; aa460 to aa467 of SEQ ID NO:7, 11, 12, or 13; aa461 to aa468 of SEQ ID NO:7, 11, 12, or 13; aa462 to aa469 of SEQ ID NO:7, 11, 12, or 13; aa463 to aa470 of SEQ ID NO:7, 11, 12, or 13; aa464 to aa471 of SEQ ID NO:7, 11, 12, or 13; aa465 to aa472 of SEQ ID NO:7, 11, 12, or 13; aa466 to aa473 of SEQ ID NO:7, 11, 12, or 13; aa467 to aa474 of SEQ ID NO:7, 11, 12, or 13; aa468 to aa475 of SEQ ID NO:7, 11, 12, or 13; aa469 to aa476 of SEQ ID NO:7, 11, 12, or 13; aa470 to aa477 of SEQ ID NO:7, 11, 12, or 13; aa471 to aa478 of SEQ ID NO:7, 11, 12, or 13; aa472 to aa479 of SEQ ID NO:7, 11, 12, or 13; aa473 to aa480 of SEQ ID NO:7, 11, 12, or 13; aa474 to aa481 of SEQ ID NO:7, 11, 12, or 13; aa475 to aa482 of SEQ ID NO:7, 11, 12, or 13; aa476 to aa483 of SEQ ID NO:7, 11, 12, or 13; aa477 to aa484 of SEQ ID NO:7, 11, 12, or 13; aa478 to aa484 of SEQ ID NO:7, 11, 12, or 13; aa479 to aa486 of SEQ ID NO:7, 11, 12, or 13; aa480 to aa487 of SEQ ID NO:7, 11, 12, or 13; aa481 to aa488 of SEQ ID NO:7, 11, 12, or 13; aa482 to aa489 of SEQ ID NO:7, 11, 12, or 13; aa483 to aa490 of SEQ ID NO:7, 11, 12, or 13; aa484 to aa491 of SEQ ID NO:7, 11, 12, or 13; aa485 to aa492 of SEQ ID NO:7, 11, 12, or 13; aa486 to aa493 of SEQ ID NO:7, 11, 12, or 13; aa487 to aa494 of SEQ ID NO:7, 11, 12, or 13; aa488 to aa495 of SEQ ID NO:7, 11, 12, or 13; aa489 to aa496 of SEQ ID NO:7, 11, 12, or 13; aa490 to aa497 of SEQ ID NO:7, 11, 12, or 13; aa491 to aa498 of SEQ ID NO:7, 11, 12, or 13; aa492 to aa499 of SEQ ID NO:7, 11, 12, or 13; aa493 to aa500 of SEQ ID NO:7, 11, 12, or 13; aa494 to aa501 of SEQ ID NO:7, 11, 12, or 13; aa495 to aa502 of SEQ ID NO:7, 11, 12, or 13; aa496 to aa503 of SEQ ID NO:7, 11, 12, or 13; aa497 to aa504 of SEQ ID NO:7, 11, 12, or 13; aa498 to aa505 of SEQ ID NO:7, 11, 12, or 13; aa499 to aa506 of SEQ ID NO:7, 11, 12, or 13; aa500 to aa507 of SEQ ID NO:7, 11, 12, or 13 aa501 to aa508 of SEQ ID NO:7, 11, 12, or 13 aa502 to aa509 of SEQ ID NO:7, 11, 12, or 13; aa503 to aa510 of SEQ ID NO:7,

11, 12, or 13; aa504 to aa511 of SEQ ID NO:7, 11, 12, or 13; aa505 to aa512 of SEQ ID NO:7, 11, 12, or 13; aa506 to aa513 of SEQ ID NO:7, 11, 12, or 13; aa507 to aa514 of SEQ ID NO:7, 11, 12, or 13; aa508 to aa515 of SEQ ID NO:7, 11, 12, or 13; aa509 to aa516 of SEQ ID NO:7, 11, 12, or 13; aa510 to aa517 of SEQ ID NO:7, 11, 12, or 13; aa511 to aa518 of SEQ ID NO:7, 11, 12, or 13; aa512 to aa519 of SEQ ID NO:7, 11, 12, or 13; aa513 to aa520 of SEQ ID NO:7, 11, 12, or 13; aa514 to aa521 of SEQ ID NO:7, 11, 12, or 13; aa515 to aa522 of SEQ ID NO:7, 11, 12, or 13; aa516 to aa523 of SEQ ID NO:7, 11, 12, or 13; aa517 to aa524 of SEQ ID NO:7, 11, 12, or 13; aa518 to aa525 of SEQ ID NO:7, 11, 12, or 13; aa519 to aa526 of SEQ ID NO:7, 11, 12, or 13; aa520 to aa527 of SEQ ID NO:7, 11, 12, or 13; aa521 to aa528 of SEQ ID NO:7, 11, 12, or 13; aa522 to aa529 of SEQ ID NO:7, 11, 12, or 13; aa523 to aa530 of SEQ ID NO:7, 11, 12, or 13; aa524 to aa531 of SEQ ID NO:7, 11, 12, or 13; aa525 to aa532 of SEQ ID NO:7, 11, 12, or 13; aa526 to aa533 of SEQ ID NO:7, 11, 12, or 13; aa527 to aa534 of SEQ ID NO:7, 11, 12, or 13; aa528 to aa535 of SEQ ID NO:7, 11, 12, or 13; aa529 to aa536 of SEQ ID NO:7, 11, 12, or 13; aa530 to aa537 of SEQ ID NO:7, 11, 12, or 13; aa531 to aa538 of SEQ ID NO:7, 11, 12, or 13; aa532 to aa539 of SEQ ID NO:7, 11, 12, or 13; aa533 to aa540 of SEQ ID NO:7, 11, 12, or 13; aa534 to aa541 of SEQ ID NO:7, 11, 12, or 13; aa535 to aa542 of SEQ ID NO:7, 11, 12, or 13; aa536 to aa543 of SEQ ID NO:7, 11, 12, or 13; aa537 to aa544 of SEQ ID NO:7, 11, 12, or 13; aa538 to aa545 of SEQ ID NO:7, 11, 12, or 13; aa539 to aa546 of SEQ ID NO:7, 11, 12, or 13; aa540 to aa547 of SEQ ID NO:7, 11, 12, or 13; aa541 to aa548 of SEQ ID NO:7, 11, 12, or 13; aa542 to aa549 of SEQ ID NO:7, 11, 12, or 13; aa543 to aa550 of SEQ ID NO:7, 11, 12, or 13; aa544 to aa551 of SEQ ID NO:7, 11, 12, or 13; aa545 to aa552 of SEQ ID NO:7, 11, 12, or 13; aa546 to aa553 of SEQ ID NO:7, 11, 12, or 13; aa547 to aa554 of SEQ ID NO:7, 11, 12, or 13; aa548 to aa555 of SEQ ID NO:7, 11, 12, or 13; aa549 to aa556 of SEQ ID NO:7, 11, 12, or 13; aa550 to aa557 of SEQ ID NO:7, 11, 12, or 13; aa551 to aa558 of SEQ ID NO:7, 11, 12, or 13; aa552 to aa559 of SEQ ID NO:7, 11, 12, or 13; aa553 to aa560 of SEQ ID NO:7, 11, 12, or 13; aa554 to aa561 of SEQ ID NO:7, 11, 12, or 13; aa555 to aa562 of SEQ ID NO:7, 11, 12, or 13; aa556 to aa563 of SEQ ID NO:7, 11, 12, or 13; aa557 to aa564 of SEQ ID NO:7, 11,

12, or 13; aa558 to aa565 of SEQ ID NO:7, 11, 12, or 13; aa559 to aa566 of SEQ ID NO:7, 11, 12, or 13; aa560 to aa567 of SEQ ID NO:7, 11, 12, or 13; aa561 to aa568 of SEQ ID NO:7, 11, 12, or 13; aa562 to aa569 of SEQ ID NO:7, 11, 12, or 13; aa563 to aa570 of SEQ ID NO:7, 11, 12, or 13; aa564 to aa571 of SEQ ID NO:7, 11, 12, or 13; aa565 to aa572 of SEQ ID NO:7, 11, 12, or 13; aa566 to aa573 of SEQ ID NO:7, 11, 12, or 13; aa567 to aa574 of SEQ ID NO:7, 11, 12, or 13; aa568 to aa575 of SEQ ID NO:7, 11, 12, or 13; aa569 to aa576 of SEQ ID NO:7, 11, 12, or 13; aa570 to aa577 of SEQ ID NO:7, 11, 12, or 13; aa571 to aa578 of SEQ ID NO:7, 11, 12, or 13; aa572 to aa579 of SEQ ID NO:7, 11, 12, or 13; aa573 to aa580 of SEQ ID NO:7, 11, 12, or 13; aa574 to aa581 of SEQ ID NO:7, 11, 12, or 13; aa575 to aa582 of SEQ ID NO:7, 11, 12, or 13; aa576 to aa583 of SEQ ID NO:7, 11, 12, or 13; aa577 to aa584 of SEQ ID NO:7, 11, 12, or 13; aa578 to aa585 of SEQ ID NO:7, 11, 12, or 13; aa579 to aa586 of SEQ ID NO:7, 11, 12, or 13; aa580 to aa587 of SEQ ID NO:7, 11, 12, or 13; aa581 to aa588 of SEQ ID NO:7, 11, 12, or 13; aa582 to aa589 of SEQ ID NO:7, 11, 12, or 13; aa583 to aa590 of SEQ ID NO:7, 11, 12, or 13; aa584 to aa591 of SEQ ID NO:7, 11, 12, or 13; aa585 to aa592 of SEQ ID NO:7, 11, 12, or 13; aa586 to aa593 of SEQ ID NO:7, 11, 12, or 13; aa587 to aa594 of SEQ ID NO:7, 11, 12, or 13; aa588 to aa595 of SEQ ID NO:7, 11, 12, or 13; aa589 to aa596 of SEQ ID NO:7, 11, 12, or 13; aa590 to aa597 of SEQ ID NO:7, 11, 12, or 13; aa591 to aa598 of SEQ ID NO:7, 11, 12, or 13; aa592 to aa599 of SEQ ID NO:7, 11, 12, or 13; aa593 to aa600 of SEQ ID NO:7, 11, 12, or 13; aa594 to aa601 of SEQ ID NO:7, 11, 12, or 13; aa595 to aa602 of SEQ ID NO:7, 11, 12, or 13; aa596 to aa603 of SEQ ID NO:7, 11, 12, or 13; aa597 to aa604 of SEQ ID NO:7, 11, 12, or 13; aa598 to aa605 of SEQ ID NO:7, 11, 12, or 13; aa599 to aa606 of SEQ ID NO:7, 11, 12, or 13; aa600 to aa607 of SEQ ID NO:7, 11, 12, or 13; aa601 to aa608 of SEQ ID NO:7, 11, 12, or 13; aa602 to aa609 of SEQ ID NO:7, 11, 12, or 13; aa603 to aa610 of SEQ ID NO:7, 11, 12, or 13; aa604 to aa611 of SEQ ID NO:7, 11, 12, or 13; aa605 to aa612 of SEQ ID NO:7, 11, 12, or 13; aa606 to aa613 of SEQ ID NO:7, 11, 12, or 13; aa607 to aa614 of SEQ ID NO:7, 11, 12, or 13; aa608 to aa615 of SEQ ID NO:7, 11, 12, or 13; aa609 to aa616 of SEQ ID NO:7, 11, 12, or 13; aa610 to aa617 of SEQ ID NO:7, 11, 12, or 13; aa611 to aa618 of SEQ ID NO:7, 11,

12, or 13; aa612 to aa619 of SEQ ID NO:7, 11, 12, or 13; aa613 to aa620 of SEQ ID NO:7, 11, 12, or 13; aa614 to aa621 of SEQ ID NO:7, 11, 12, or 13; aa615 to aa622 of SEQ ID NO:7, 11, 12, or 13; aa616 to aa623 of SEQ ID NO:7, 11, 12, or 13; aa617 to aa624 of SEQ ID NO:7, 11, 12, or 13; aa618 to aa625 of SEQ ID NO:7, 11, 12, or 13; aa619 to aa626 of SEQ ID NO:7, 11, 12, or 13; aa620 to aa627 of SEQ ID NO: 12 or 13; aa621 to aa628 of SEQ ID NO: 12 or 13; aa622 to aa629 of SEQ ID NO:12 or 13; aa623 to aa630 of SEQ ID NO:12 or 13; aa624 to aa631 of SEQ ID NO:12 or 13; aa625 to aa632 of SEQ ID NO:12 or 13; aa626 to aa633 of SEQ ID NO:12 or 13; aa627 to aa634 of SEQ ID NO:12 or 13; aa628 to aa635 of SEQ ID NO:12 or 13; aa629 to aa636 of SEQ ID NO:12 or 13; aa630 to aa637 of SEQ ID NO:12 or 13; aa631 to aa638 of SEQ ID NO:12 or 13; aa632 to aa639 of SEQ ID NO:12 or 13; aa633 to aa640 of SEQ ID NO:12 or 13; aa634 to aa641 of SEQ ID NO:12 or 13; aa635 to aa642 of SEQ ID NO:12 or 13; aa636 to aa643 of SEQ ID NO:12 or 13; aa637 to aa644 of SEQ ID NO:12 or 13; aa638 to aa645 of SEQ ID NO:12 or 13; aa639 to aa646 of SEQ ID NO:12 or 13; aa640 to aa647 of SEQ ID NO:12 or 13; aa641 to aa648 of SEQ ID NO:12 or 13; aa642 to aa649 of SEQ ID NO:12 or 13; aa643 to aa650 of SEQ ID NO:12 or 13; aa644 to aa651 of SEQ ID NO:12 or 13; aa645 to aa652 of SEQ ID NO:12 or 13; aa646 to aa653 of SEQ ID NO:12 or 13; aa647 to aa654 of SEQ ID NO:12 or 13; aa648 to aa655 of SEQ ID NO:12 or 13; aa649 to aa656 of SEQ ID NO:12 or 13; aa650 to aa657 of SEQ ID NO:12 or 13; aa651 to aa658 of SEQ ID NO:12 or 13; aa652 to aa659 of SEQ ID NO:12 or 13; aa653 to aa660 of SEQ ID NO:12 or 13; aa654 to aa661 of SEQ ID NO:12 or 13; aa655 to aa662 of SEQ ID NO:12 or 13; aa656 to aa663 of SEQ ID NO:12 or 13; aa657 to aa664 of SEQ ID NO:12 or 13; aa658 to aa665 of SEQ ID NO:12 or 13; aa659 to aa666 of SEQ ID NO:12 or 13; aa660 to aa667 of SEQ ID NO:12 or 13; aa661 to aa668 of SEQ ID NO:12 or 13; aa662 to aa669 of SEQ ID NO:12 or 13; aa663 to aa670 of SEQ ID NO:12 or 13; aa664 to aa671 of SEQ ID NO:12 or 13; aa665 to aa672 of SEQ ID NO:12 or 13; aa666 to aa673 of SEQ ID NO:12 or 13; aa667 to aa674 of SEQ ID NO:12 or 13; aa668 to aa675 of SEQ ID NO:12 or 13; aa669 to aa676 of SEQ ID NO:12 or 13; aa670 to aa677 of SEQ ID NO:12 or 13; aa671 to aa678 of SEQ ID

NO:12 or 13; aa672 to aa679 of SEQ ID NO:12 or 13; aa673 to aa680 of SEQ ID NO:12 or 13; aa674 to aa681 of SEQ ID NO:12 or 13; aa675 to aa682 of SEQ ID NO:12 or 13; aa676 to aa683 of SEQ ID NO:12 or 13; aa677 to aa684 of SEQ ID NO:12 or 13; aa678 to aa684 of SEQ ID NO:12 or 13; aa679 to aa686 of SEQ ID NO:12 or 13; aa680 to aa687 of SEQ ID NO:12 or 13; aa681 to aa688 of SEQ ID NO:12 or 13; aa682 to aa689 of SEQ ID NO:12 or 13; aa683 to aa690 of SEQ ID NO:12 or 13; aa684 to aa691 of SEQ ID NO:12 or 13; aa685 to aa692 of SEQ ID NO:12 or 13; aa686 to aa693 of SEQ ID NO:12 or 13; aa687 to aa694 of SEQ ID NO:12 or 13; aa688 to aa695 of SEQ ID NO:12 or 13; aa689 to aa696 of SEQ ID NO:12 or 13; aa690 to aa697 of SEQ ID NO:12 or 13; aa691 to aa698 of SEQ ID NO:12 or 13; aa692 to aa699 of SEQ ID NO:12 or 13; aa693 to aa700 of SEQ ID NO:12 or 13; aa694 to aa701 of SEQ ID NO:12 or 13; aa695 to aa702 of SEQ ID NO:12 or 13; aa696 to aa703 of SEQ ID NO:12 or 13; aa697 to aa704 of SEQ ID NO:12 or 13; aa698 to aa705 of SEQ ID NO:12 or 13; aa699 to aa706 of SEQ ID NO:12 or 13; aa700 to aa707 of SEQ ID NO:12 or 13; aa701 to aa708 of SEQ ID NO:12 or 13; aa702 to aa709 of SEQ ID NO:12 or 13; aa703 to aa710 of SEQ ID NO:12 or 13; aa704 to aa711 of SEQ ID NO:12 or 13; aa705 to aa712 of SEQ ID NO:12 or 13; aa76 to aa713 of SEQ ID NO:12 or 13; aa707 to aa714 of SEQ ID NO:12 or 13; aa708 to aa715 of SEQ ID NO:12 or 13; aa709 to aa716 of SEQ ID NO:12 or 13; aa710 to aa717 of SEQ ID NO:12 or 13; aa711 to aa718 of SEQ ID NO:12 or 13; aa712 to aa719 of SEQ ID NO:12 or 13; aa713 to aa720 of SEQ ID NO:12 or 13; aa714 to aa721 of SEQ ID NO:12 or 13; aa715 to aa722 of SEQ ID NO:12 or 13; aa716 to aa723 of SEQ ID NO:12 or 13; aa717 to aa724 of SEQ ID NO:12 or 13; aa718 to aa725 of SEQ ID NO:12 or 13; aa719 to aa726 of SEQ ID NO:12 or 13; aa720 to aa727 of SEQ ID NO:12 or 13; aa721 to aa728 of SEQ ID NO:12 or 13; aa722 to aa729 of SEQ ID NO:12 or 13; aa723 to aa730 of SEQ ID NO:12 or 13; aa724 to aa731 of SEQ ID NO:12 or 13; aa725 to aa732 of SEQ ID NO:12 or 13; aa726 to aa733 of SEQ ID NO:12 or 13; aa727 to aa734 of SEQ ID NO:12 or 13; aa728 to aa735 of SEQ ID NO:12 or 13; aa729 to aa736 of SEQ ID NO:12 or 13; aa730 to aa737 of SEQ ID NO:12 or 13; aa731 to aa738 of SEQ ID NO:12 or 13; aa732 to aa739 of SEQ ID

NO:12 or 13; aa733 to aa740 of SEQ ID NO:12 or 13; aa734 to aa741 of SEQ ID NO:12 or 13; aa735 to aa742 of SEQ ID NO:12 or 13; aa736 to aa743 of SEQ ID NO:12 or 13; aa737 to aa744 of SEQ ID NO:12 or 13; aa738 to aa745 of SEQ ID NO:12 or 13; aa739 to aa746 of SEQ ID NO:12 or 13; aa740 to aa747 of SEQ ID NO:12 or 13; aa741 to aa748 of SEQ ID NO:12 or 13; aa742 to aa749 of SEQ ID NO:12 or 13; aa743 to aa750 of SEQ ID NO:12 or 13; aa744 to aa751 of SEQ ID NO:12 or 13; aa745 to aa752 of SEQ ID NO:12 or 13; aa746 to aa753 of SEQ ID NO:12 or 13; aa747 to aa754 of SEQ ID NO:12 or 13; aa748 to aa755 of SEQ ID NO:12 or 13; aa749 to aa756 of SEQ ID NO:12 or 13; aa750 to aa757 of SEQ ID NO:12 or 13; aa751 to aa758 of SEQ ID NO:12 or 13; aa752 to aa759 of SEQ ID NO:12 or 13; aa753 to aa760 of SEQ ID NO:12 or 13; aa754 to aa761 of SEQ ID NO:12 or 13; aa755 to aa762 of SEQ ID NO:12 or 13; aa756 to aa763 of SEQ ID NO:12 or 13; aa757 to aa764 of SEQ ID NO:12 or 13; aa758 to aa765 of SEQ ID NO:12 or 13; aa759 to aa766 of SEQ ID NO:12 or 13; aa760 to aa767 of SEQ ID NO:12 or 13; aa761 to aa768 of SEQ ID NO:12 or 13; aa762 to aa769 of SEQ ID NO:12 or 13; aa763 to aa770 of SEQ ID NO:12 or 13; aa764 to aa771 of SEQ ID NO:12 or 13; aa765 to aa772 of SEQ ID NO:12 or 13; aa766 to aa773 of SEQ ID NO:12 or 13; aa767 to aa774 of SEQ ID NO:12 or 13; aa768 to aa775 of SEQ ID NO:12 or 13; aa769 to aa776 of SEQ ID NO:12 or 13; aa770 to aa777 of SEQ ID NO:12 or 13; aa771 to aa778 of SEQ ID NO:12 or 13; aa772 to aa779 of SEQ ID NO:12 or 13; aa773 to aa780 of SEQ ID NO:12 or 13; aa774 to aa781 of SEQ ID NO:12 or 13; aa775 to aa782 of SEQ ID NO:12 or 13; aa776 to aa783 of SEQ ID NO:12 or 13; aa777 to aa784 of SEQ ID NO:12 or 13; aa778 to aa784 of SEQ ID NO:12 or 13; aa779 to aa786 of SEQ ID NO:12 or 13; aa780 to aa787 of SEQ ID NO:12 or 13; aa781 to aa788 of SEQ ID NO:12 or 13; aa782 to aa789 of SEQ ID NO:12 or 13; aa783 to aa790 of SEQ ID NO:12 or 13; aa784 to aa791 of SEQ ID NO:12 or 13; aa785 to aa792 of SEQ ID NO:12 or 13; aa786 to aa793 of SEQ ID NO:12 or 13; aa787 to aa794 of SEQ ID NO:12 or 13; aa788 to aa795 of SEQ ID NO:12 or 13; aa789 to aa796 of SEQ ID NO:12 or 13; aa790 to aa797 of SEQ ID NO:12 or 13; aa791 to aa798 of SEQ ID NO:12 or 13; aa792 to aa799 of SEQ ID

NO:12 or 13; aa793 to aa800 of SEQ ID NO:12 or 13; aa794 to aa801 of SEQ ID NO:12 or 13; aa795 to aa802 of SEQ ID NO:12 or 13; aa796 to aa803 of SEQ ID NO:12 or 13; aa797 to aa804 of SEQ ID NO:12 or 13; aa798 to aa805 of SEQ ID NO:12 or 13; aa799 to aa806 of SEQ ID NO:12 or 13; aa800 to aa807 of SEQ ID NO:12 or 13; aa801 to aa808 of SEQ ID NO:12 or 13; aa802 to aa809 of SEQ ID NO:12 or 13; aa803 to aa810 of SEQ ID NO:12 or 13; aa804 to aa811 of SEQ ID NO:12 or 13; aa805 to aa812 of SEQ ID NO:12 or 13; aa806 to aa813 of SEQ ID NO:12 or 13; aa807 to aa814 of SEQ ID NO:12 or 13; aa808 to aa815 of SEQ ID NO:12 or 13; aa809 to aa816 of SEQ ID NO:12 or 13; aa810 to aa817 of SEQ ID NO:12 or 13; aa811 to aa818 of SEQ ID NO:12 or 13; aa812 to aa819 of SEQ ID NO:12 or 13; aa813 to aa820 of SEQ ID NO:12 or 13; aa814 to aa821 of SEQ ID NO:12 or 13; aa815 to aa822 of SEQ ID NO:12 or 13; aa816 to aa823 of SEQ ID NO:12 or 13; aa817 to aa824 of SEQ ID NO:12 or 13; aa818 to aa825 of SEQ ID NO:12 or 13; aa819 to aa826 of SEQ ID NO:12 or 13; aa820 to aa827 of SEQ ID NO:12 or 13; aa821 to aa828 of SEQ ID NO:12 or 13; aa822 to aa829 of SEQ ID NO:12 or 13; aa823 to aa830 of SEQ ID NO:12 or 13; aa824 to aa831 of SEQ ID NO:12 or 13; aa825 to aa832 of SEQ ID NO:12 or 13; aa826 to aa833 of SEQ ID NO:12 or 13; aa827 to aa834 of SEQ ID NO:12 or 13; aa828 to aa835 of SEQ ID NO:12 or 13; aa829 to aa836 of SEQ ID NO:12 or 13; aa830 to aa837 of SEQ ID NO:12 or 13; aa831 to aa838 of SEQ ID NO:12 or 13; aa832 to aa839 of SEQ ID NO:12 or 13; aa833 to aa840 of SEQ ID NO:12 or 13; aa834 to aa841 of SEQ ID NO:12 or 13; aa835 to aa842 of SEQ ID NO:12 or 13; aa836 to aa843 of SEQ ID NO:12 or 13; aa837 to aa844 of SEQ ID NO:12 or 13; aa838 to aa845 of SEQ ID NO:12 or 13; aa839 to aa846 of SEQ ID NO:12 or 13; aa840 to aa847 of SEQ ID NO:12 or 13; aa841 to aa848 of SEQ ID NO:12 or 13; aa842 to aa849 of SEQ ID NO:12 or 13; aa843 to aa850 of SEQ ID NO:12 or 13; aa844 to aa851 of SEQ ID NO:12 or 13; aa845 to aa852 of SEQ ID NO:12 or 13; aa846 to aa853 of SEQ ID NO:12 or 13; aa847 to aa854 of SEQ ID NO:12 or 13; aa848 to aa855 of SEQ ID NO:12 or 13; aa849 to aa856 of SEQ ID NO:12 or 13; aa850 to aa857 of SEQ ID NO:12 or 13; aa851 to aa858 of SEQ ID NO:12 or 13; aa852 to aa859 of SEQ ID

NO:12 or 13; aa853 to aa860 of SEQ ID NO:12 or 13; aa854 to aa861 of SEQ ID NO:12 or 13; aa855 to aa862 of SEQ ID NO:12 or 13; aa856 to aa863 of SEQ ID NO:12 or 13; aa857 to aa864 of SEQ ID NO:12 or 13; aa858 to aa865 of SEQ ID NO:12 or 13; aa859 to aa866 of SEQ ID NO:12 or 13; aa860 to aa867 of SEQ ID NO:12 or 13; aa861 to aa868 of SEQ ID NO:12 or 13; aa862 to aa869 of SEQ ID NO:12 or 13; aa863 to aa870 of SEQ ID NO:12 or 13; aa864 to aa871 of SEQ ID NO:12 or 13; aa865 to aa872 of SEQ ID NO:12 or 13; aa866 to aa873 of SEQ ID NO:12 or 13; aa867 to aa874 of SEQ ID NO:12 or 13; aa868 to aa875 of SEQ ID NO:12 or 13; aa869 to aa876 of SEQ ID NO:12 or 13; aa870 to aa877 of SEQ ID NO:12 or 13; aa871 to aa878 of SEQ ID NO:12 or 13; aa872 to aa879 of SEQ ID NO:12 or 13; aa873 to aa880 of SEQ ID NO:12 or 13; aa874 to aa881 of SEQ ID NO:12 or 13; aa875 to aa882 of SEQ ID NO:12 or 13; aa876 to aa883 of SEQ ID NO:12 or 13; aa877 to aa884 of SEQ ID NO:12 or 13; aa878 to aa884 of SEQ ID NO:12 or 13; aa879 to aa886 of SEQ ID NO:12 or 13; aa880 to aa887 of SEQ ID NO:12 or 13; aa881 to aa888 of SEQ ID NO:12 or 13; aa882 to aa889 of SEQ ID NO:12 or 13; aa883 to aa890 of SEQ ID NO:12 or 13; aa884 to aa891 of SEQ ID NO:12 or 13; aa885 to aa892 of SEQ ID NO:12 or 13; aa886 to aa893 of SEQ ID NO:12 or 13; aa887 to aa894 of SEQ ID NO:12 or 13; aa888 to aa895 of SEQ ID NO:12 or 13; aa889 to aa896 of SEQ ID NO:12 or 13; aa890 to aa897 of SEQ ID NO:12 or 13; aa891 to aa898 of SEQ ID NO:12 or 13; aa892 to aa899 of SEQ ID NO:12 or 13; aa893 to aa900 of SEQ ID NO:12 or 13; aa894 to aa901 of SEQ ID NO:12 or 13; aa895 to aa902 of SEQ ID NO:12 or 13; aa896 to aa903 of SEQ ID NO:12 or 13; aa897 to aa904 of SEQ ID NO:12 or 13; aa898 to aa905 of SEQ ID NO:12 or 13; aa899 to aa906 of SEQ ID NO:12 or 13; aa900 to aa907 of SEQ ID NO:12 or 13; aa901 to aa908 of SEQ ID NO:12 or 13; aa902 to aa909 of SEQ ID NO:12 or 13; aa903 to aa910 of SEQ ID NO:12 or 13; aa904 to aa911 of SEQ ID NO:12 or 13; aa905 to aa912 of SEQ ID NO:12 or 13; aa906 to aa913 of SEQ ID NO:12 or 13; aa907 to aa914 of SEQ ID NO:12 or 13; aa908 to aa915 of SEQ ID NO:12 or 13; aa909 to aa916 of SEQ ID NO:12 or 13; aa910 to aa917 of SEQ ID NO:12 or 13; aa911 to aa918 of SEQ ID NO:12 or 13; aa912 to aa919 of SEQ ID

NO:12 or 13; aa913 to aa920 of SEQ ID NO:12 or 13; aa914 to aa921 of SEQ ID NO:12 or 13; aa915 to aa922 of SEQ ID NO:12 or 13; aa916 to aa923 of SEQ ID NO:12 or 13; aa917 to aa924 of SEQ ID NO:12 or 13; aa918 to aa925 of SEQ ID NO:12 or 13; aa919 to aa926 of SEQ ID NO:12 or 13; aa920 to aa927 of SEQ ID NO:12 or 13; aa921 to aa928 of SEQ ID NO:12 or 13; aa922 to aa929 of SEQ ID NO:12 or 13; aa923 to aa930 of SEQ ID NO:12 or 13; aa924 to aa931 of SEQ ID NO:12 or 13; aa925 to aa932 of SEQ ID NO:12 or 13; aa926 to aa933 of SEQ ID NO:12 or 13; aa927 to aa934 of SEQ ID NO:12 or 13; aa928 to aa935 of SEQ ID NO:12 or 13; aa929 to aa936 of SEQ ID NO:12 or 13; aa930 to aa937 of SEQ ID NO:12 or 13; aa931 to aa938 of SEQ ID NO:12 or 13; aa932 to aa939 of SEQ ID NO:12 or 13; aa933 to aa940 of SEQ ID NO:12 or 13; aa934 to aa941 of SEQ ID NO:12 or 13; aa935 to aa942 of SEQ ID NO:12 or 13; aa936 to aa943 of SEQ ID NO:12 or 13; aa937 to aa944 of SEQ ID NO:12 or 13; aa938 to aa945 of SEQ ID NO:12 or 13; aa939 to aa946 of SEQ ID NO:12 or 13; aa940 to aa947 of SEQ ID NO:12 or 13; aa941 to aa948 of SEQ ID NO:12 or 13; aa942 to aa949 of SEQ ID NO:12 or 13; aa943 to aa950 of SEQ ID NO:12 or 13; aa944 to aa951 of SEQ ID NO:12 or 13; aa945 to aa952 of SEQ ID NO:12 or 13; aa946 to aa953 of SEQ ID NO:12 or 13; aa947 to aa954 of SEQ ID NO:12 or 13; aa948 to aa955 of SEQ ID NO:12 or 13; aa949 to aa956 of SEQ ID NO:12 or 13; aa950 to aa957 of SEQ ID NO:12 or 13; aa951 to aa958 of SEQ ID NO:12 or 13; aa952 to aa959 of SEQ ID NO:12 or 13; aa953 to aa960 of SEQ ID NO:12 or 13; aa954 to aa961 of SEQ ID NO:12 or 13; aa955 to aa962 of SEQ ID NO:12 or 13; aa956 to aa963 of SEQ ID NO:12 or 13; aa957 to aa964 of SEQ ID NO:12 or 13; aa958 to aa965 of SEQ ID NO:12 or 13; aa959 to aa966 of SEQ ID NO:12 or 13; aa960 to aa967 of SEQ ID NO:12 or 13; aa961 to aa968 of SEQ ID NO:12 or 13; aa962 to aa969 of SEQ ID NO:12 or 13; aa963 to aa970 of SEQ ID NO:12 or 13; aa964 to aa971 of SEQ ID NO:12 or 13; aa965 to aa972 of SEQ ID NO:12 or 13; aa966 to aa973 of SEQ ID NO:12 or 13; aa967 to aa974 of SEQ ID NO:12 or 13; aa968 to aa975 of SEQ ID NO:12 or 13; aa969 to aa976 of SEQ ID NO:12 or 13; aa970 to aa977 of SEQ ID NO:12 or 13; aa971 to aa978 of SEQ ID NO:12 or 13; aa972 to aa979 of SEQ ID

NO:12 or 13; aa973 to aa980 of SEQ ID NO:12 or 13; aa974 to aa981 of SEQ ID NO:12 or 13; aa975 to aa982 of SEQ ID NO:12 or 13; aa976 to aa983 of SEQ ID NO:12 or 13; aa977 to aa984 of SEQ ID NO:12 or 13; aa978 to aa984 of SEQ ID NO:12 or 13; aa979 to aa986 of SEQ ID NO:12 or 13; aa980 to aa987 of SEQ ID NO:12 or 13; aa981 to aa988 of SEQ ID NO:12 or 13; aa982 to aa989 of SEQ ID NO:12 or 13; aa983 to aa990 of SEQ ID NO:12 or 13; aa984 to aa991 of SEQ ID NO:12 or 13; aa985 to aa992 of SEQ ID NO:12 or 13; aa986 to aa993 of SEQ ID NO:12 or 13; aa987 to aa994 of SEQ ID NO:12 or 13; aa988 to aa995 of SEQ ID NO:12 or 13; aa989 to aa996 of SEQ ID NO:12 or 13; aa990 to aa997 of SEQ ID NO:12 or 13; aa991 to aa998 of SEQ ID NO:12 or 13; aa992 to aa999 of SEQ ID NO:12 or 13; aa993 to aa1000 of SEQ ID NO:12 or 13; aa994 to aa1001 of SEQ ID NO:12 or 13; aa995 to aa1002 of SEQ ID NO:12 or 13; aa996 to aa1003 of SEQ ID NO:12 or 13; aa997 to aa1004 of SEQ ID NO:12 or 13; aa998 to aa1005 of SEQ ID NO:12 or 13; aa999 to aa1006 of SEQ ID NO:12 or 13; aa1000 to aa1007 of SEQ ID NO:12 or 13; aa1001 to aa1008 of SEQ ID NO:12 or 13; aa1002 to aa1009 of SEQ ID NO:12 or 13; aa1003 to aa1010 of SEQ ID NO:12 or 13; aa1004 to aa1011 of SEQ ID NO:12 or 13; aa1005 to aa1012 of SEQ ID NO:12 or 13; aa1006 to aa1013 of SEQ ID NO:12 or 13; aa1007 to aa1014 of SEQ ID NO:12 or 13; aa1008 to aa1015 of SEQ ID NO:12 or 13; aa1009 to aa1016 of SEQ ID NO:12 or 13; aa1010 to aa1017 of SEQ ID NO:12 or 13; aa1011 to aa1018 of SEQ ID NO:12 or 13; aa1012 to aa1019 of SEQ ID NO:12 or 13; aa1013 to aa1020 of SEQ ID NO:12 or 13; aa1014 to aa1021 of SEQ ID NO:12 or 13; aa1015 to aa1022 of SEQ ID NO:12 or 13; aa1016 to aa1023 of SEQ ID NO:12 or 13; aa1017 to aa1024 of SEQ ID NO:12 or 13; aa1018 to aa1025 of SEQ ID NO:12 or 13; aa1019 to aa1026 of SEQ ID NO:12 or 13; aa1020 to aa1027 of SEQ ID NO:12 or 13; aa1021 to aa1028 of SEQ ID NO:12 or 13; aa1022 to aa1029 of SEQ ID NO:12 or 13; aa1023 to aa1030 of SEQ ID NO:12 or 13; aa1024 to aa1031 of SEQ ID NO:12 or 13; aa1025 to aa1032 of SEQ ID NO:12 or 13; aa1026 to aa1033 of SEQ ID NO:12 or 13; aa1027 to aa1034 of SEQ ID NO:12 or 13; aa1028 to aa1035 of SEQ ID NO:12 or 13; aa1029 to aa1036 of SEQ ID NO:12 or 13; aa1030 to aa1037 of SEQ ID NO:12 or 13; aa1031 to aa1038 of SEQ ID NO:12 or 13; aa1032 to aa1039 of SEQ ID

NO:12 or 13; aa1033 to aa1040 of SEQ ID NO:12 or 13; aa1034 to aa1041 of SEQ ID NO:12 or 13; aa1035 to aa1042 of SEQ ID NO:12 or 13; aa1036 to aa1043 of SEQ ID NO:12 or 13; aa1037 to aa1044 of SEQ ID NO:12 or 13; aa1038 to aa1045 of SEQ ID NO:12 or 13; aa1039 to aa1046 of SEQ ID NO:12 or 13; aa1040 to aa1047 of SEQ ID NO:12 or 13; aa1041 to aa1048 of SEQ ID NO:12 or 13; aa1042 to aa1049 of SEQ ID NO:12 or 13; aa1043 to aa1050 of SEQ ID NO:12 or 13; aa1044 to aa1051 of SEQ ID NO:12 or 13; aa1045 to aa1052 of SEQ ID NO:12 or 13; aa1046 to aa1053 of SEQ ID NO:12 or 13; aa1047 to aa1054 of SEQ ID NO:12 or 13; aa1048 to aa1055 of SEQ ID NO:13; aa1049 to aa1056 of SEQ ID NO: 13; aa1050 to aa1057 of SEQ ID NO:13; aa1051 to aa1058 of SEQ ID NO: 13; aa1052 to aa1059 of SEQ ID NO:13; aa1053 to aa1060 of SEQ ID NO:13; aa1054 to aa1061 of SEQ ID NO:13; aa1055 to aa1062 of SEQ ID NO:13; aa1056 to aa1063 of SEQ ID NO:13; aa1057 to aa1064 of SEQ ID NO:13; aa1058 to aa1065 of SEQ ID NO:13; aa1059 to aa1066 of SEQ ID NO:13; aa1060 to aa1067 of SEQ ID NO:13; aa1061 to aa1068 of SEQ ID NO:13; aa1062 to aa1069 of SEQ ID NO:13; aa1063 to aa1070 of SEQ ID NO:13; aa1064 to aa1071 of SEQ ID NO:13; aa1065 to aa1072 of SEQ ID NO:13; aa1066 to aa1073 of SEQ ID NO:13; aa1067 to aa1074 of SEQ ID NO:13; aa1068 to aa1075 of SEQ ID NO:13; aa1069 to aa1076 of SEQ ID NO:13; aa1070 to aa1077 of SEQ ID NO:13; and aa1071 to aa1078 of SEQ ID NO:13. The coding sequence of fragments can be constructed by cleaving the unwanted nucleotides from the mutant or native TIAM2 protein coding sequences.

Fusions are fragment, mutant, or native TIAM2 proteins with additional amino acids at either or both of the termini. The additional amino acid sequence generally is not homologous to sequence found in native TIAM2 polypeptides. The additional amino acid residues can facilitate expression, detection, or activity of the polypeptide, for example. The additional amino acid sequence can also be used as linker to construct multimers of TIAM2 proteins. All fusion polypeptides exhibit the desired sequence homology, immunological or biological activity. Recombinant TIAM2 fusion proteins can be produced using the preferred cell system, baculovirus.

After inserting the TIAM2 DNA sequence into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to those skilled in the art and fully described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987).

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above-described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate construct (transfer vector).

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 base pairs downstream from the ATT; *see* Luckow and Summers, *Virology 17*:31 (1989).

The plasmid usually also contains the polyhedron polyadenylation signal (Miller *et al. Ann. Rev. Microbiol.*, *42*:177 (1988)) and a prokaryotic ampicillin-resistance (*amp*) gene and origin of replication for selection and propagation in *E. coli*.

Baculovirus transfer vectors usually contain a baculovirus promoter. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen *et al.*, (1986) "The

Regulation of Baculovirus Gene Expression," in: The Molecular Biology of Baculoviruses (ed. Walter Doerfler); EPO Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak *et al.*, *J. Gen. Virol. 69*:765 (1988).

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell *et al. Gene*, *73*:409 (1988)). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human α-interferon, Maeda *et al.*, *Nature 315*:592 (1985); human gastrin-releasing peptide, Lebacq-Verheyden *et al.*, *Molec. Cell. Biol. 8*:3129 (1988); human IL-2, Smith *et al.*, *Proc. Nat'l Acad. Sci. USA, 82*:8404 (1985); mouse IL-3, (Miyajima *et al.*, *Gene 58*:273 (1987); and human glucocerebrosidase, Martin *et al. DNA 7*:99 (1988), can also be used to provide for secretion in insects.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus, usually by co-transfection. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See Summers and Smith; Ju *et al.* (1987); Smith *et al.*, *Mol. Cell. Biol. 3*:2156 (1983); and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller *et al.*, *Bioessays 4*:91 (1989).

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Methods to identify recombinant viruses are described in "Current Protocols in Microbiology" Vol. 2 (Ausubel *et al.* eds) at 16.8 (Supp. 10, 1990); Summers and Smith; Miller *et al.* (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, inter alia: *Aedes aegypti* , *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (PCT Pub. No. WO 89/046699; Carbonell *et al.*, *J. Virol. 56*:153 (1985); Wright *Nature 321*:718 (1986); Smith *et al.*, *Mol. Cell. Biol. 3*:2156 (1983); and see generally, Fraser, *et al.* In *Vitro Cell. Dev. Biol. 25*:225 (1989)).

The TIAM2 protein can be used in an assay for inhibitors and for preparing antibodies directed to TIAM2. TIAM2 protein may also be useful as a factor that promotes the growth of cancer cells in culture. The TIAM2 protein may be combined with the pharmaceutically acceptable carrier noted above for use with the TIAM2 antisense molecule.

The present invention will now be illustrated by reference to the following examples which set forth particularly advantageous embodiments. However, it should be noted that these embodiments are illustrative and are not to be construed as restricting the invention in any way.

EXAMPLES

EXAMPLE 1:

IDENTIFICATION OF THE TIAM2 GENE

In a degenerative PCR-based screen for new genes, a single 377-nucleotide clone (DP-75) was identified, with high homology to the mouse and human TIAM1 genes. Due to its similarity to TIAM1, this new gene and its predicted protein product were designated T-cell lymphoma invasion and metastasis 2 (TIAM2).

Northern analysis with the DP-75 probe revealed a ~3.3-kb transcript in cerebral cortex, occipital pole, frontal lobe, and temporal lobe, and a ~4.4-kb transcript in testis and cerebellum (Figures 1A and 1B). Through cDNA library screens and RACE cloning, overlapping clones were isolated and compiled into full-length

sequences representing the ~4.4- and ~3.3-kb messages. To perform Northern analyses, Human Multiple Tissue Northern Blots (Clontech) were hybridized with DP-75 probe in SpeedHyb solution (BIOS lab) at 65° C. Blots were washed two times with 0.5 x SSC, 0.5% SDS at room temperature for 10 min. and one time with 0.1 x SSC, 0.5% SDS at 65° C for I hr.

The compiled cDNA for the ~4.4-kb message is 4586 nucleotides long, without polyadenylation, and is predicted to encode the 1077-amino-acid-long form of the TIAM2 protein ($TIAM2_L$). The first ATG is at nucleotide 51 and is in good context for translation. An alignment of the predicted protein to TIAM1 (Figure 3) shows significant identity beginning with the first methionine of TIAM2 and a methionine at position 576 of TIAM1, a region that corresponds to the last 14 amino acids of the coiled coil domain of TIAM1 (aa 548-590). The identity continues into the EX domain, which extends from amino acid 599 to amino acid 691 of human TIAM1. A comparison of the TIAM1/$TIAM2_L$ EX domains shows 62% overall identity, but with 85% identity in a 60-aa subdomain. Following the EX domain, alternative splicing of a 72-nucleotide miniexon leads to the presence or absence of a 24-amino acid sequence with no significant identity to TIAM1 or *SIF* sequences. Following the alternatively spliced region is a sequence with weak identity to TIAM1 (18%), but which contains the conserved residues of PDZ or DHR domains (Michiels et al., Nature 375:338-340, 1995). The remaining sequence contains identity to the DH catalytic region (70%) and the carboxyl-terminal PH domain (54%).

The compiled sequence for the ~3.3-kb message is 3344 nucleotides long and is predicted to encode a short form (626 aa) of TIAM2 ($TIAM2_S$). The ~3.3-kb message differs from the ~4.4-kb message in that the first 104 nucleotides are unique, suggesting that the ~3.3-kb message results from initiation at a promoter distinct from that of the ~4.4-kb message. The initiator methionine at position 162 coincides with a methionine at amino acid 428 of the 4.4-kb TIAM2 predicted protein and for the remainder of the protein shows complete identity with $TIAM2_L$. Similar to what is seen with the TIAM1 gene (Habets et al., Cell 77:537-549, 1994), there are two upstream ATGs in the ~3.3-kb message that precede the long open reading frame (nucleotide

162), but only the start codon at nucleotide 162 is in good context for translation. A comparison with TIAM1 shows that $TIAM2_S$, begins immediately upstream (24 aa) of the DH domain.

Northern analysis with a probe isolated from the 5' end of the ~4.4-kb message identifies only the larger TIAM2 transcript (Figure 1C), while a probe isolated from the 3' end of TIAM2 identifies both transcripts (Figure 1D).

A Radiation Hybrid Mapping experiment was performed as follows. TIAM2-specific oligos AGTCCTACCTCATCAAGCCG (SEQ ID NO:14) and TTAGTGCTTCCGTCAGGTGG (SEQ ID NO:15) were synthesized and used with the G3 panel (Research Genetics) as per the manufacturer's instructions. The resulting PCR products were analyzed by gel electrophoresis, and the resulting data were analyzed according to methods of Stewart and Cox, *In* "Genome Mapping: A Practical Approach" (P. Dear et al. Ed.) pp. 73-93, Oxford Univ. Press, Oxford, 1997. A two-point maximum-likelihood analysis showed that the markers were linked to D6S1556 (6q25) with a lod score of 1000.

EXAMPLE 2

IN SITU HYBRIDIZATION

Oligos designed from nucleotides 566-586 (AGTACAGAAGGTT-CTGGAGCG) and from nucleotides 1348-1367 (TACCAGGCGATCCTTACACG) of the human 3.3 kb TIAM2 cDNA were used to perform PCR on mouse brain cDNA (Clontech). Thirty cycles of PCR performed with an annealing temperature of 40° C led to the production of a product (MDP75) of the expected size (802 nt). Cloning and sequencing confirmed that MDP75 was 90% identical to human TIAM2 throughout the length of the clone. Mouse embryos (CD-1) were processed for whole-mount RNA *in situ* hybridization following the protocol of Nieto *et al.*, *In* "Methods in Avian Embryology" (M. Bronner-Fraser, Ed.), Academic Press, San Diego (1995). Adult brain sections were processed using the protocol for frozen sections as described (Schaeren-Wiemers and Gerfin-Moser, *Histochemistry 100*:431-440 (1993)).

The expression of TIAM2 mRNA in the embryonic day 13.5 (E13.5) and adult mouse brain was examined by *in situ* hybridization with a digoxigenin-labeled ribo-probe. Discrete and reproducible labeling was observed at both ages with an antisense probe; no labeling was detected with a sense probe. In both E13.5 and adult brain, the cellular localization of TIAM-2 transcripts is consistent with TIAM2 expression in neurons rather than glial cells. At E13.5, TIAM2 transcripts are present throughout the telencephalon (Figure 2A). Parasagittal sections through the telencephalon demonstrate that expression is localized to the pial surface where early born postmitotic neurons are located (Figure 2B). TIAM2 is not expressed in the proliferating neural precursor cells that line the ventricles.

In the adult mouse brain, TIAM2 continues to be expressed primarily in regions derived from the telencephalon (Figures 2C-2G, Table 1). Areas of strong expression include the olfactory bulb, cerebral cortex, caudate putamen, and hippocampus. The hippocampus expression is highly specific in that cells of the dentate gyrus and regions derived from it, induseum griseum and fasciola cinerea; and pyramidal cells of one subregion of the CA fields (CA2) are labeled very strongly while CA1 and CA3 show very little expression. TIAM2 is also strongly expressed in the ependymal cells of the lateral surface of the lateral ventricles (Figure 2D). This region is a germinal zone where neurons are generated throughout adult life. The majority of neurons born in this region are believed to migrate to the olfactory bulb.

TABLE 1

EXPRESSION OF TIAM2 MRNA IN THE ADULT MOUSE BRAIN

Olfactory bulb	
Mitral cell layer	+++
Glomerular layer	++
Granular cell layer	+++
Olfactory nuclei	++
Cerebral cortex	+++
Hippocampus	
CA1	+
CA2	++++
CA3	+
Dentate gyrus	++++
Insidium gresium	++++
Fasciola cinerea	++++
Amygdala	+
Septum	++
Caudate putamen	+++
Ependymal and subependymal layers	+++
Thalamus	
Reticular nucleus	+++
Ventral posterolateral nucleus	+++
Pons	++

Note. Relative levels of expression based on bright-field microscopy. The highest levels and lowest levels of expression are indicated by ++++ and +, respectively.

EXAMPLE 3

GDP-GTP EXCHANGE ACTIVITY

An amino-terminal HIS tagged version of the 3.3-kb $TIAM_S$ protein (His-TIAM2) was expressed in a baculovirus system. A protein of the predicted molecular mass (75,661 Da) was partially purified on a nickel column, and its identity was confirmed by amino-terminal sequencing and mass spectrometry. To generate the His-tagged $TIAM2_S$ expression construct, a linker-adapter was produced such that it added an *Eco*RI site to the 5' end of the first 29 bp of the $TIAM2_S$ sequence (GAATTCATGGAAGGACCGCGGGA-GAATCAGGATCC). This kinased linker/adapter, terminating in a *Bam*HI site (bases 23-29 of the $TIAM2_S$ sequence), was ligated to a cDNA clone DP-75#1, which had been linearized by a *Bam*HI fragment that encompassed the entire coding region of $TIAM2_S$. A subsequent *Sal*I digest liberated an ~2000-bp *Eco*RI-*Sal*I fragment that encompassed the entire coding region of $TIAM2_S$. The fragment was isolated, purified, and ligated into *Eco*RI-*Sal*I-digested pBlueBacHis2B vector (Invitrogen). The resulting clone (pHIS- $TIAM2_S$) was confirmed by sequence analysis.

pHIS-$TIAM2_S$ was introduced into SF9 cells, and clones expressing high levels of pHIS- $TIAM2_S$ were identified by Western blot. Following clonal expansion, pHIS- $TIAM2_S$-expressing cells were concentrated, lysed, and applied to a nickel column (Invitrogen). pHIS-$TIAM2_S$ protein was purified from the column as per the manufacturer's recommendations except that following the 50 mM imidazole wash, the protein was directly eluted in 500 mM imidazole. The fractions containing pHIS-$TIAM2_S$ were concentrated on a Centricon (Amicon) 30 column spun at 5000*g* for 30 min.

His-$TIAM2_S$ was tested for its ability to stimulate the GDP-GTP exchange activity with Rac or Ras proteins and was compared to the activity of SOS, a known Ras GEF (Jefferson et al., Oncogene 16:2303-2310, 1998). In these experiments (Figure 4), immunoprecipitated Rac or Ras protein was preloaded with GDP and incubated in [^{32}P]GTP-containing buffer alone or in buffer containing either His-

$TIAM2_S$ or purified SOS. In multiple experiments, His-$TIAM_2$ showed approximately fivefold stimulation of Rac exchange activity at 30 min relative to buffer alone, while SOS showed a modest stimulation (1.6×) at best (Figure 4A). Conversely, when His-$TIAM2_S$ and SOS were incubated with RAS, SOS demonstrated a four- to fivefold stimulation of exchange activity while His-$TIAM2_S$ had little or no effect (Figure 4B).

EXAMPLE 4:

UTILIZING POLYNUCLEOTIDES TO DIAGNOSE CANCER

In a dot blot assay, TIAM2 was hybridized to RNA from both cancerous and normal tissue. The source of cancerous tissue include renal, thyroid, breast, colon, ureter, lung, nose, stomach, esophagus, liver, lymphoma, uterus, bladder, rectum, and brain.

The blots were from BioChain Institute, Inc., San Leandro, California, USA. ExpressHyb™ hybridization buffer (Clontech, Palo Alto, California, USA) was used for the blotting at 68°C with TIAM2 (SEQ ID NO:1) at $1x10^6$ cpm/ml for 2 hours.

As shown in Figure 5, in four of four thyroid samples, SEQ ID NO:1 mRNA levels were higher in the cancer than the normal samples. In two of the four colon samples, SEQ ID NO:1 mRNA levels were higher in the cancer than the normal samples. In one of two ureter samples, the SEQ ID NO:1 (TIAM2) mRNA level were higher in the cancer than the normal sample. In one of four breast samples, the SEQ ID NO:1 mRNA level were higher in the cancer than the normal sample. In one of four renal sample, the SEQ ID NO:1 mRNA level were higher in the cancer than the normal sample. In all other tissue types tested, the SEQ ID NO:1 mRNA levels were the same or higher in the normal samples than the cancer samples.

SEQ ID NO:1

AAGGCCTTTGTTGGGTGCCCGGAACCCCACCAAGCAGCATTCCTCACGCTG
GAGTCCTACCTCATCAAGCCGGTTCAGAGAGCGCTCAGGTACCCGCTGCTG
CTCAAGGAGCTGGTGTCCCTGACGGACCAGGAGAGCGAGGAGCACTACCAC
CTGACGGAAGCACTAAAGGCAATGGAGAAAGTAGCGAGCCACATCAATGA
GATGCAGAAGATCTATGAGGATTATGGGACCGTGTTTGACCAGCTAGTAGC
TGAGCAGAGCGGAACAGAGAAGGAGGTAACAGAACTTTCGATGGGAGAGC
TTCTGATGCACTCTACGGTTTCCTGGTTGAACCCAATGTTGATCCCCGGGG

EXAMPLE 5:

ISOLATION OF SEQ ID NO:6

SEQ ID NO:6 was isolated from a frontal cortex library utilizing a phage vector Stratagene, La Jolla, California, USA). The library was probed with SEQ ID NO:1, which was generated by a random primed label with a final radioactive count of approximately $1x10^6$ cpm/ml. The probe was labeled according to manufacturer's instruction with a RediPrime™ DNA labeling kit (Amersham, Arlington Heights, Illinois, USA).

The phage library was propagated and plated onto twenty plates according to the manufacturer's instructions with a 3.0-5.0 x 10^5 plaques/plate. The plaques were transferred to a nitrocellulose membranes. Each membrane was incubated with the SEQ ID NO:1 probe for 2 hours at 65°C in ExpressHyb™ hybridization solution purchased from Clontech, Palo Alto, California, USA. The filters were washed according to the Clontech instruction. Film was exposed to the membranes to identify putative positive plaques containing the desired TIAM2 polynucleotide.

A second round of plating and hybridization was performed to identify a single positive plaque. The positive plaques from the first round were propagated and

plated onto agar medium according to the instructions provided by Stratagene. The plaques were transferred to filters. These filters were incubated with the SEQ ID NO:1 probe. The probe and hybridization conditions were the same as described above. Positive plaques were identified and propagated.

According to manufacturer's instructions, a BlueScript plasmid was rescued from the phage vector. The EcoRI insert from the plasmid was sequenced. The polynucleotide sequence is shown in SEQ ID NO:6.

The present invention has been described with reference to specific embodiments. However, this application is intended to cover those changes and substitutions which may be made by those skilled in the art without departing from the spirit and the scope of the appended claims.

Deposit Information:

The following materials were deposited with the American Type Culture Collection:

Name	Deposit Date	Accession No.
Escherichia coli INVαF' DP 75	25 April 1996	98030

The above materials have been deposited with the American Type Culture Collection, Manassas, Virginia, under the accession numbers indicated. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for purposes of Patent Procedure. The deposits will be maintained for a period of 30 years following issuance of this patent, or for the enforceable life of the patent, whichever is greater. Upon issuance of the patent, the deposits will be available to the public from the ATCC without restriction.

These deposits are provided merely as convenience to those of skill in the art, and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained within the deposited materials, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by

reference and are controlling in the event of any conflict with the written description of sequences herein. A license may be required to make, use, or sell the deposited materials, and no such license is granted hereby.

CLAIMS

WHAT IS CLAIMED:

1. An isolated nucleic acid molecule comprising a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding amino acids from about 1 to about 626 of SEQ ID NO:11;

(b) a polynucleotide encoding amino acids from about 2 to about 626 of SEQ ID NO:11;

(c) a polynucleotide encoding amino acids from about 1 to about 626 of SEQ ID NO:12;

(d) a polynucleotide encoding amino acids from about 2 to about 626 of SEQ ID NO:12;

(e) a polynucleotide encoding amino acids from about 1 to about 626 of SEQ ID NO:13;

(f) a polynucleotide encoding amino acids from about 2 to about 626 of SEQ ID NO:13;

(g) a polynucleotide at least 80 % identical to the polynucleotide of (a), (b), (c), (d), (e) or (f); and

(h) the polynucleotide complement of the polynucleotide of (a), (b), (c), (d), (e), (f), (g).

2. An isolated nucleic acid molecule comprising 10 contiguous nucleotides from the coding region of SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

3. The isolated nucleic acid molecule of claim 2, which comprises 20 contiguous nucleotides from the coding region of SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

4. The isolated nucleic acid molecule of claim 3, which comprises

50 contiguous nucleotides from the coding region of SEQ ID NO:8, SEQ ID NO:9 or SEQ ID NO:10.

5. An isolated nucleic acid molecule comprising a polynucleotide encoding a polypeptide wherein, except for at least one conservative amino acid substitution, said polypeptide has an amino acid sequence selected from the group consisting of:

(a) a polynucleotide encoding amino acids from about 1 to about 626 of SEQ ID NO:11;

(b) a polynucleotide encoding amino acids from about 2 to about 626 of SEQ ID NO:11;

(c) a polynucleotide encoding amino acids from about 1 to about 626 of SEQ ID NO:12;

(d) a polynucleotide encoding amino acids from about 2 to about 626 of SEQ ID NO:12;

(e) a polynucleotide encoding amino acids from about 1 to about 626 of SEQ ID NO:13; and

(f) a polynucleotide encoding amino acids from about 2 to about 626 of SEQ ID NO:13.

6. A method of making a recombinant vector comprising inserting a nucleic acid molecule of claim 1 into a vector in operable linkage to a promoter.

7. A recombinant vector produced by the method of claim 6.

8. A method of making a recombinant host cell comprising introducing the recombinant vector of claim 7 into a host cell.

9. A recombinant host cell produced by the method of claim 8.

10. A recombinant method of producing a polypeptide, comprising culturing the recombinant host cell of claim 9 under conditions such that said polypeptide is expressed and recovering said polypeptide.

11. An isolated polypeptide comprising amino acids at least 90% identical to amino acids selected from the group consisting of:

(a) amino acids from about 1 to about 626 of SEQ ID NO:11;

(b) amino acids from about 2 to about 626 of SEQ ID NO:11;

(c) amino acids from about 1 to about 626 of SEQ ID NO:12;

(d) amino acids from about 2 to about 626 of SEQ ID NO:12;

(e) amino acids from about 1 to about 626 of SEQ ID NO:13; and

(f) amino acids from about 2 to about 626 of SEQ ID NO:13.

12. An isolated polypeptide wherein, except for at least one conservative amino acid substitution, said polypeptide has an amino acid sequence selected from the group consisting of:

(a) amino acids from about 1 to about 626 of SEQ ID NO:11;

(b) amino acids from about 2 to about 626 of SEQ ID NO:11;

(c) amino acids from about 1 to about 626 of SEQ ID NO:12;

(d) amino acids from about 2 to about 626 of SEQ ID NO:12;

(e) amino acids from about 1 to about 626 of SEQ ID NO:13; and

(f) amino acids from about 2 to about 626 of SEQ ID NO:13.

13. An isolated polypeptide comprising amino acids selected from the group consisting of:

(a) amino acids from about 1 to about 626 of SEQ ID NO:11;

(b) amino acids from about 2 to about 626 of SEQ ID NO:11;

(c) amino acids from about 1 to about 626 of SEQ ID NO:12;

(d) amino acids from about 2 to about 626 of SEQ ID NO:12;

(e) amino acids from about 1 to about 626 of SEQ ID NO:13; and

(f) amino acids from about 2 to about 626 of SEQ ID NO:13.

14. An epitope-bearing portion of the polypeptide of SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13.

15. The epitope-bearing portion of claim 14, which comprises 10 contiguous amino acids of SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13.

16. The epitope-bearing portion of claim 15, which comprises 20 contiguous amino acids of SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13.

17. An isolated antibody that binds specifically to the polypeptide of claim 11.

18. An isolated antibody that binds specifically to a polypeptide of claim 12.

19. An isolated antibody that binds specifically to the polypeptide of claim 13.

20. An antisense vector comprising (i) an antisense polynucleotide that comprises a sequence capable of hybridizing to at least one of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof under stringent conditions wherein said polynucleotide is capable of hybridizing to the mRNA of native human TIAM2; and (ii) and polynucleotide comprising a sequence capable of initiating transcription of said antisense polynucleotide.

21. The antisense vector of claim 20, wherein said sequences to initiate transcription are derived from a retrovirus, an adenovirus, or an adeno-associated virus.

22. The antisense vector of claim 21, further comprising an origin of replication.

23. A method of inhibiting expression of TIAM2, comprising:

(a) providing an antisense polynucleotide comprising a nucleic acid sequence capable of hybridizing to at least one of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof under stringent conditions and is capable of hybridizing to native human TIAM2 mRNA; and

(b) contacting said antisense polynucleotide with said TIAM2.

24. A method for detecting hyperproliferative cells in a sample comprising

(a) providing a probe polynucleotide comprising a sequence capable of hybridizing to at least one of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof under stringent conditions;

(b) contacting the probe with the polynucleotides of the sample cell under conditions permitting formation of polynucleotide hybrids; and

(c) detecting the hybrids.

25. A method to detect TIAM2 polypeptides in a sample comprising:

(a) providing an antibody that specifically binds to a TIAM2 polypeptide, wherein said TIAM2 polypeptide comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence capable of hybridizing to at least one of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof;

(b) contacting said antibody to the sample under conditions permitting the formation of antibody/antigen complexes; and

(c) detecting the complexes.

26. A method of inhibiting the replication of a cell comprising

(a) providing an antibody that specifically binds to a TIAM2 polypeptide, wherein said TIAM2 polypeptide comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence capable of hybridizing to at least one of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 or fragments thereof; and

(b) contacting said antibody to the sample under conditions permitting the formation of antibody/antigen complexes.

Figure 1

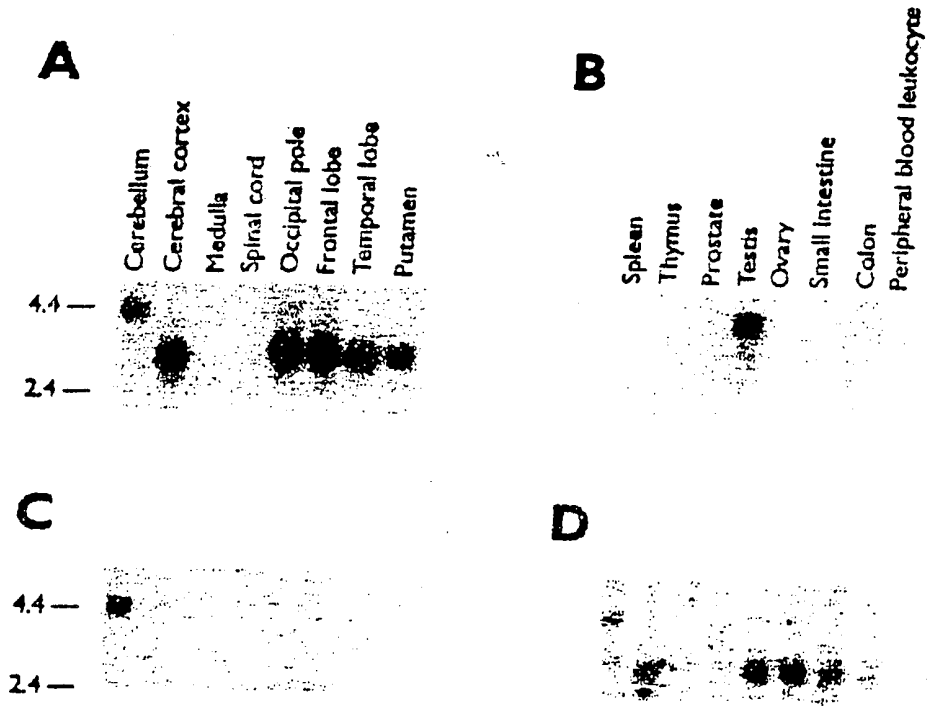

A
Cerebellum
Cerebral cortex
Medulla
Spinal cord
Occipital pole
Frontal lobe
Temporal lobe
Putamen
4.4
2.4
B
Spleen
Thymus
Prostate
Testis
Ovary
Small intestine
Colon
Peripheral blood leukocyte
C
4.4
2.4
D

Figure 2

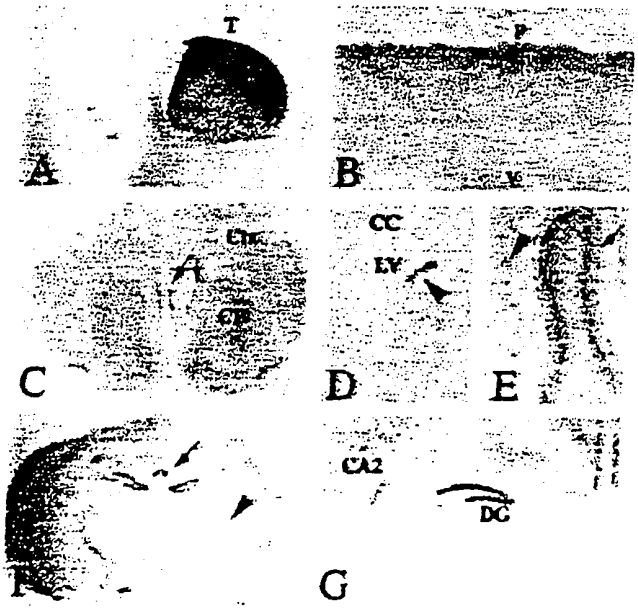


Figure 4

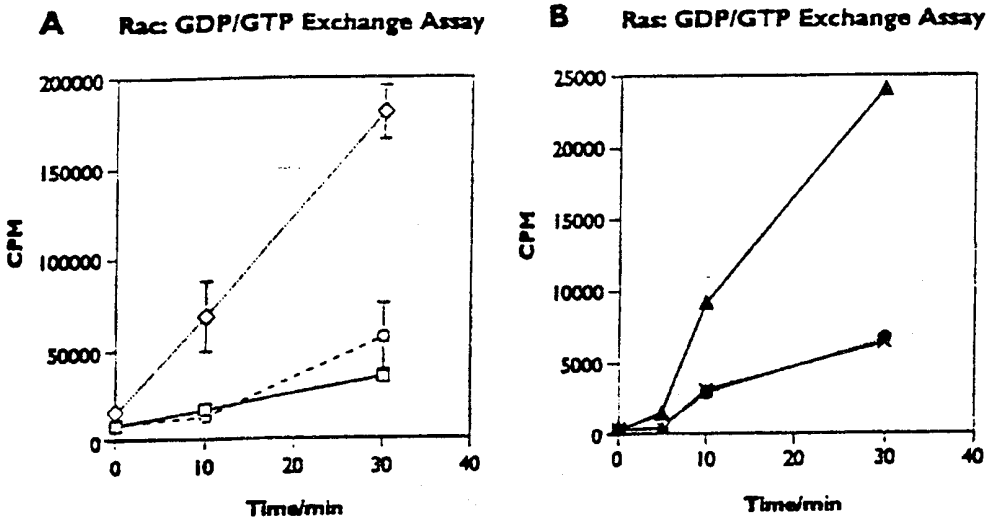
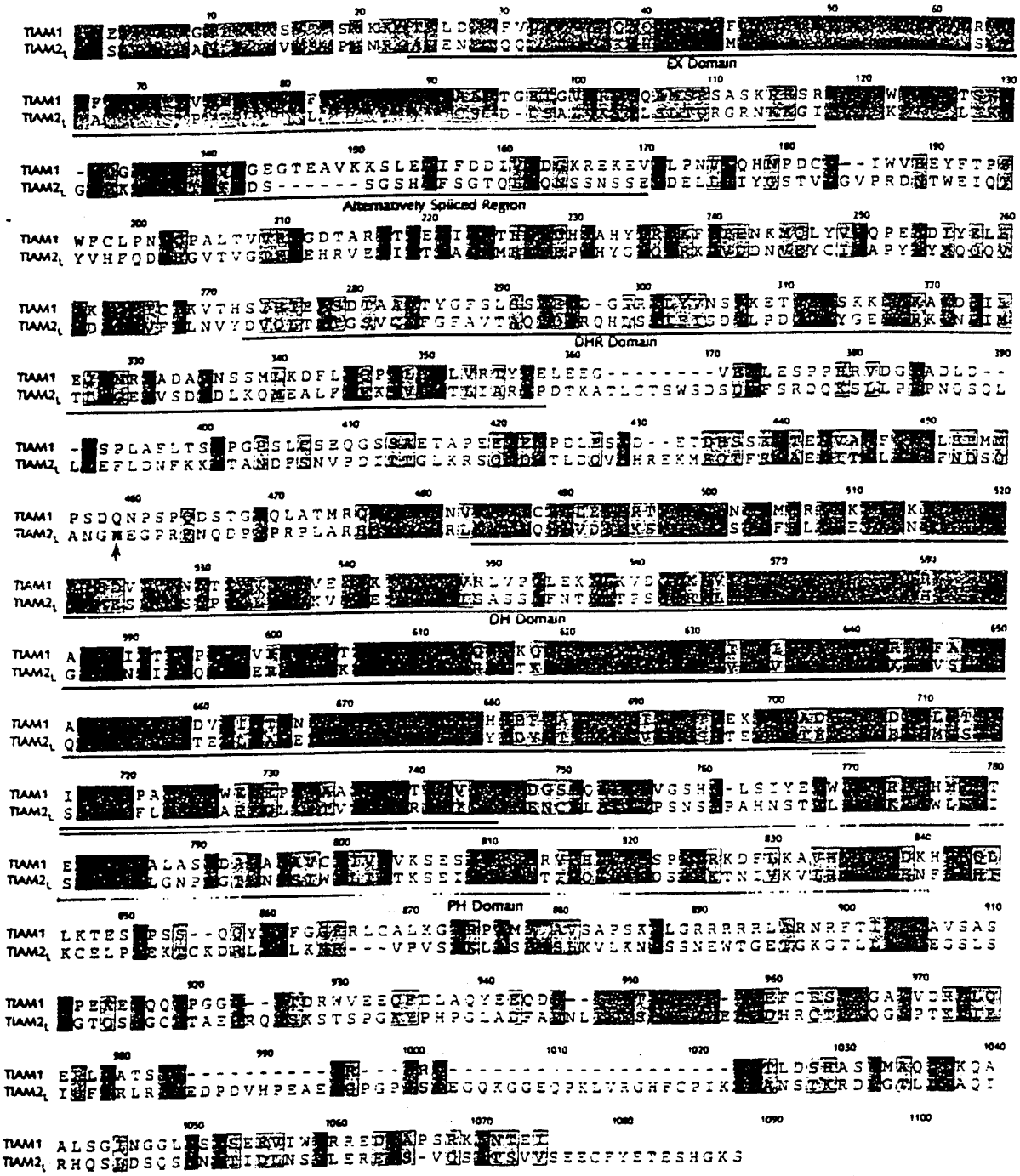


Figure 3


TIAM1
TIAM2L
EX Domain
Alternatively Spliced Region
DHR Domain
DH Domain
PH Domain

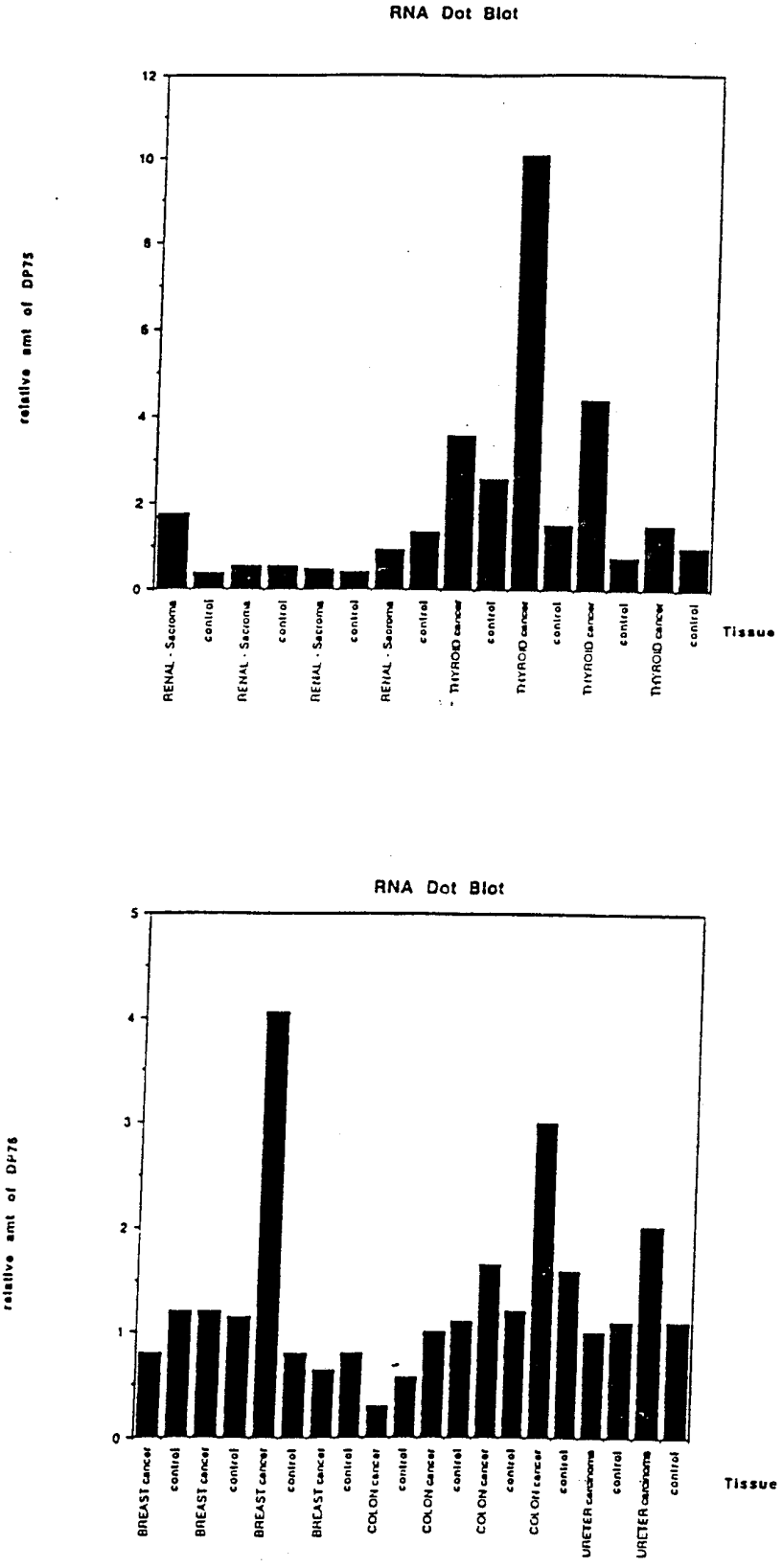

RNA Dot Blot
relative amt of DP75
12
10
8
6
4
2
0
RENAL - Sarcoma
control
RENAL - Sarcoma
control
RENAL - Sarcoma
control
RENAL - Sarcoma
control
THYROID cancer
control
THYROID cancer
control
THYROID cancer
control
THYROID cancer
control
Tissue
RNA Dot Blot
relative amt of DP76
5
4
3
2
1
0
BREAST cancer
control
BREAST cancer
control
BREAST cancer
control
BREAST cancer
control
COLON cancer
control
COLON cancer
control
COLON cancer
control
COLON cancer
control
URETER carcinoma
control
URETER carcinoma
control
Tissue

Figure 5

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Duhl, David

(ii) TITLE OF INVENTION: DNA Encoding T-Cell Lymphoma Invasion and Metastasis 2 and a Process for its Use

(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Chiron Corporation
(B) STREET: 4560 Horton Street
(C) CITY: Emeryville
(D) STATE: California
(E) COUNTRY: United States
(F) ZIP: 94608

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

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Potter, Jane E. R.
(B) REGISTRATION NUMBER: 33,332
(C) REFERENCE/DOCKET NUMBER: 1203.200PC

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 206-622-4900
(B) TELEFAX: 206-682-6031

(2) INFORMATION For SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

AAGGCCTTTG TTGGGTGCCC GGAACCCCAC CAAGCAGCAT TCCTCACGCT GGAGTCCTAC
60

CTCATCAAGC CGGTTCAGAG AGCGCTCAGG TACCCGCTGC TGCTCAAGGA GCTGGTGTCC
120

CTGACGGACC AGGAGAGCGA GGAGCACTAC CACCTGACGG AAGCACTAAA GGCAATGGAG
180

AAAGTAGCGA GCCACATCAA TGAGATGCAG AAGATCTATG AGGATTATGG GACCGTGTTT
240

GACCAGCTAG TAGCTGAGCA GAGCGGAACA GAGAAGGAGG TAACAGAACT TTCGATGGGA
300

GAGCTTCTGA TGCACTCTAC GGTTTCCTGG TTGAACCCAA TGTTGATCCC CGGGG
355

(2) INFORMATION For SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

AAGGCCTTTG TTGGGTGCC
19

(2) INFORMATION For SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

AAGGCCTTTG YTGGNYNCC
19

(2) INFORMATION For SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

CCCCGGGGAT CAACATTGGG TT
22

(2) INFORMATION For SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

CCCCGGGGAT VADVADDGGR TT
22

(2) INFORMATION For SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2089 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

GAATTCCACT TAGATGTGAT CCGTTCTCCC AGAGGGAGCA GGTTTCTTTG AACTTTTCCT
60

TTTTATGTAC AGCATAGTGC TGAGCAGATC ACTGCACTGT GCAGGAGTTT TAACGACAGT
120

CAGGCCAACG GCATGGAAGG ACCGCGGGAG AATCAGGATC CTCCTCCGAG GCCTCTGGCC
180

CGCCACCTGT CTGATGCAGA CCGCCTCCGC AAAGTCATCC AGGAGCTTGT GGACACAGAG
240

AAGTCCTACG TGAAGGATTT GAGCTGCCTC TTTGAATTAT ACTTGGAGCC ACTTCAGAAT
300

GAGACCTTTC TTACCCAAGA TGAGATGGAG TCACTTTTTG GAAGTTTGCC AGAGATGCTT
360

GAGTTTCAGA AGGTGTTTCT GGAGACCCTG GAGGATGGGA TTTCAGCATC ATCTGACTTT
420

AACACCCTAG AAACCCCCTC ACAGTTTAGA AAATTACTGT TTTCCCTTGG AGGCTCTTTC
480

CTTTATTACG CGGACCACTT TAAACTGTAC AGTGGATTCT GTGCTAACCA TATCAAAGTA
540

CAGAAGGTTC TGGAGCGAGC TAAAACTGAC AAAGCCTTCA AGGCTTTTCT GGACGCCCGG
600

AACCCCACCA AGCAGCATTC CTCCACGCTG GAGTCCTACC TCATCAAGCC GGTTCAGAGA
660

GTGCTCAAGT ACCCGCTGCT GCTCAAGGAG CTGGTGTCCC TGACGGACCA GGAGAGCGAG
720

GAGCACTACC ACCTGACGGA AGCACTAAAG GCAATGGAGA AAGTAGCGAG CCACATCAAT
780

GAGATGCAGA AGATCTATGA GGATTATGGG ACCGTGTTTG ACCAGCTAGT AGCTGAGCAG
840

AGCGGAACAG AGAAGGAGGT AACAGAACTT TCGATGGGAG AGCTTCTGAT GCACTCTACG
900

GTTTCCTGGT TGAATCCATT TCTGTCTCTA GGAAAAGCTA GAAAGGACCT TGAGCTCACA
960

GTATTTGTTT TTAAGAGAGC CGTCATACTG GTTTATAAAG AAAACTGCAA ACTGAAAAAG
1020

AAATTGCCCT CGAATTCCCG GCCTGCACAC AACTCTACTG ACTTGGACCC ATTTAAATTC
1080

CGCTGGTTGA TCCCCATCTC CGCGCTTCAA GTCAGACTGG GGAATCCAGC AGGGACAGAA
1140

AATAATTCCA TATGGGAACT GATCCATACG AAGTCAGAAA TAGAAGGACG GCCAGAAACC
1200

ATCTTTCAGT TGTGTTGCAG TGACAGTGAA AGCAAAACCA ACATTGTTAA GGTGATTCGT
1260

TCTATTCTGA GGGAGAACTT CAGGCGTCAC ATAAAGTGTG AATTACCACT GGAGAAAACG
1320

TGTAAGGATC GCCTGGTACC TCTTAAGAAC CGAGTTCCTG TTTCGGCCAA ATTAGCTTCA
1380

TCCAGGTCTT TAAAAGTCCT GAAGAATTCC TCCAGCAACG AGTGGACCGG TGAGACTGGC
1440

AAGGGAACCT TGCTGGACTC TGACGAGGGC AGCTTGAGCA GCGGCACCCA GAGCAGCGGC
1500

TGCCCCACGG CTGAGGGCAG GCAGGACTCC AAGAGCACTT CTCCCGGGAA ATACCCACAC
1560

CCCGGCTTGG CAGATTTTGC TGACAATCTC ATCAAAGAGA GTGACATCCT GAGCGATGAA
1620

GATGATGACC ACCGTCAGAC TGTGAAGCAG GGCAGCCCTA CTAAAGACAT CGAAATTCAG
1680

TTCCAGAGAC TGAGGATTTC CGAGGACCCA GACGTTCACC CCGAGGCTGA GCAGCAGCCT
1740

GGCCCGGAGT CGGGTGAGGG TCAGAAAGGA GGAGAGCAGC CCAAACTGGT CCGGGGGCAC
1800

TTCTGCCCCA TTAAACGAAA AGCCAACAGC ACCAAGAGGG ACAGAGGAAC TTTGCTCAAG
1860

GCGCAGATCC GTCACCAGTC CCTTGACAGT CAGTCTGAAA ATGCCACCAT CGACCTAAAT
1920

TCTGTTCTAG AGCGAGAATT CAGTGTCCAG AGTTTAACAT CTGTTGTCAG TGAGGAGTGT
1980

TTTTATGAAA CAGAGAGCCA CGGAAAATCA TAGTATGATT CAATCCAGAT ATGGGTTAAA
2040

TTCCTCATTT TACTTTTAAA CTGGTGGTAA AGTGGAAATT GCGGAATTC
2089

(2) INFORMATION For SEQ ID NO:7, 11, 12, or 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 626 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7, 11, 12, or 13:

Met Glu Gly Pro Arg Glu Asn Gln Asp Pro Pro Pro Arg Pro Leu Ala
1 5 10 15

Arg His Leu Ser Asp Ala Asp Arg Leu Arg Lys Val Ile Gln Glu Leu
20 25 30

Val Asp Thr Glu Lys Ser Tyr Val Lys Asp Leu Ser Cys Leu Phe Glu
35 40 45

Leu Tyr Leu Glu Pro Leu Gln Asn Glu Thr Phe Leu Thr Gln Asp Glu
50 55 60

Met Glu Ser Leu Phe Gly Ser Leu Pro Glu Met Leu Glu Phe Gln Lys
65 70 75 80

Val Phe Leu Glu Thr Leu Glu Asp Gly Ile Ser Ala Ser Ser Asp Phe
85 90 95

Asn Thr Leu Glu Thr Pro Ser Gln Phe Arg Lys Leu Leu Phe Ser Leu
100 105 110

Gly Gly Ser Phe Leu Tyr Tyr Ala Asp His Phe Lys Leu Tyr Ser Gly
115 120 125

Phe Cys Ala Asn His Ile Lys Val Gln Lys Val Leu Glu Arg Ala Lys
130 135 140

Thr Asp Lys Ala Phe Lys Ala Phe Leu Asp Ala Arg Asn Pro Thr Lys
145 150 155 160

Gln His Ser Ser Thr Leu Glu Ser Tyr Leu Ile Lys Pro Val Gln Arg
165 170 175

Val Leu Lys Tyr Pro Leu Leu Leu Lys Glu Leu Val Ser Leu Thr Asp
180 185 190

Gln Glu Ser Glu Glu His Tyr His Leu Thr Glu Ala Leu Lys Ala Met
195 200 205

Glu Lys Val Ala Ser His Ile Asn Glu Met Gln Lys Ile Tyr Glu Asp
210 215 220

Tyr Gly Thr Val Phe Asp Gln Leu Val Ala Glu Gln Ser Gly Thr Glu
225 230 235 240

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

Val Ser Trp Leu Asn Pro Phe Leu Ser Leu Gly Lys Ala Arg Lys Asp
260 265 270

Leu Glu Leu Thr Val Phe Val Phe Lys Arg Ala Val Ile Leu Val Tyr
275 280 285

```
Lys Glu Asn Cys Lys Leu Lys Lys Lys Leu Pro Ser Asn Ser Arg Pro
    290                 295                 300

Ala His Asn Ser Thr Asp Leu Asp Pro Phe Lys Phe Arg Trp Leu Ile
305                 310                 315                 320

Pro Ile Ser Ala Leu Gln Val Arg Leu Gly Asn Pro Ala Gly Thr Glu
                325                 330                 335

Asn Asn Ser Ile Trp Glu Leu Ile His Thr Lys Ser Glu Ile Glu Gly
            340                 345                 350

Arg Pro Glu Thr Ile Phe Gln Leu Cys Cys Ser Asp Ser Glu Ser Lys
        355                 360                 365

Thr Asn Ile Val Lys Val Ile Arg Ser Ile Leu Arg Glu Asn Phe Arg
    370                 375                 380

Arg His Ile Lys Cys Glu Leu Pro Leu Glu Lys Thr Cys Lys Asp Arg
385                 390                 395                 400

Leu Val Pro Leu Lys Asn Arg Val Pro Val Ser Ala Lys Leu Ala Ser
                405                 410                 415

Ser Arg Ser Leu Lys Val Leu Lys Asn Ser Ser Ser Asn Glu Trp Thr
            420                 425                 430

Gly Glu Thr Gly Lys Gly Thr Leu Leu Asp Ser Asp Glu Gly Ser Leu
        435                 440                 445

Ser Ser Gly Thr Gln Ser Ser Gly Cys Pro Thr Ala Glu Gly Arg Gln
    450                 455                 460

Asp Ser Lys Ser Thr Ser Pro Gly Lys Tyr Pro His Pro Gly Leu Ala
465                 470                 475                 480

Asp Phe Ala Asp Asn Leu Ile Lys Glu Ser Asp Ile Leu Ser Asp Glu
                485                 490                 495

Asp Asp Asp His Arg Gln Thr Val Lys Gln Gly Ser Pro Thr Lys Asp
            500                 505                 510

Ile Glu Ile Gln Phe Gln Arg Leu Arg Ile Ser Glu Asp Pro Asp Val
        515                 520                 525

His Pro Glu Ala Glu Gln Gln Pro Gly Pro Glu Ser Gly Glu Gly Gln
    530                 535                 540

Lys Gly Gly Glu Gln Pro Lys Leu Val Arg Gly His Phe Cys Pro Ile
545                 550                 555                 560

Lys Arg Lys Ala Asn Ser Thr Lys Arg Asp Arg Gly Thr Leu Leu Lys
```

565 570 575

Ala Gln Ile Arg His Gln Ser Leu Asp Ser Gln Ser Glu Asn Ala Thr
580 585 590

Ile Asp Leu Asn Ser Val Leu Glu Arg Glu Phe Ser Val Gln Ser Leu
595 600 605

Thr Ser Val Val Ser Glu Glu Cys Phe Tyr Glu Thr Glu Ser His Gly
610 615 620

Lys Ser
625

SEQ ID NO:8
3.3 kb message

GAGACAACGTGTGTCTTACACATCTCCCAACAGCCGACTTAGATGTGATCCGTTCTCCCAGAGGGAGCAGGTTTCTTT
GA
ACTTTTCCTTTTTATGTACAGCATAGTGCTGAGCAGATCACTGCACTGTGCAGGAGTTTTAACGACAGTCAGGCCGAC
GG
CATGGAAGGACCGCGGGAGAATCAGGATCCTCCTCCGAGGCCTCTGGCCCGCCACCTGTCTGATGCAGACCGCCTCCG
CA
AAGTCATCCAGGAGCTTGTGGACACAGAGAAGTCCTACGTGAAGGATTTGAGCTGCCTCTTTGAATTATACTTGGAGC
CA
CTTCAGAATGAGACCTTTCTTACCCAAGATGAGATGGAGTCACTTTTTGGAAGTTTGCCAGAGATGCTTGAGTTTCAG
AA
GGTGTTTCTGGAGACCCTGGAGGATGGGATTTCAGCATCATCTGACTTTAACACCCTAGAAACCCCCTCACAGTTTAG
AA
AATTACTGTTTTCCCTTGGAGGCTCTTTCCTTTATTACGCGGACCACTTTAAACTGTACAGTGGATTCTGTGCTAACC
AT
ATCAAAGTACAGAAGGTTCTGGAGCGAGCTAAAACTGACAAAGCCTTCAAGGCTTTTCTGGACGCCCGGAACCCCACC
AA
GCAGCATTCCTCCACGCTGGAGTCCTACCTCATCAAGCCGGTTCAGAGAGTGCTCAAGTACCCGCTGCTGCTCAAGGA
GC
TGGTGTCCCTGACGGACCAGGAGAGCGAGGAGCACTACCACCTGACGGAAGCACTAAAGGCAATGGAGAAAGTAGCGA
GC
CACATCAATGAGATGCAGAAGATCTATGAGGATTATGGGACCGTGTTTGACCAGCTAGTAGCTGAGCAGAGCGGAACA
GA
GAAGGAGGTAACAGAACTTTCGATGGGAGAGCTTCTGATGCACTCTACGGTTTCCTGGTTGAATCCATTTCTGTCTCT
AG
GAAAAGCTAGAAAGGACCTTGAGCTCACAGTATTTGTTTTTAAGAGAGCCGTCATACTGGTTTATAAAGAAAACTGCA
AA
CTGAAAAAGAAATTGCCCTCGAATTCCCGGCCTGCACACAACTCTACTGACTTGGACCCATTTAAATTCCGCTGGTTG

AT
CCCCATCTCCGCGCTTCAAGTCAGACTGGGGAATCCAGCAGGGACAGAAAATAATTCCATATGGGAACTGATCCATAC
GA
AGTCAGAAATAGAAGGACGGCCAGAAACCATCTTTCAGTTGTGTTGCAGTGACAGTGAAAGCAAAACCAACATTGTTA
AG
GTGATTCGTTCTATTCTGAGGGAGAACTTCAGGCGTCACATAAAGTGTGAATTACCACTGGAGAAAACGTGTAAGGAT
CG
CCTGGTACCTCTTAAGAACCGAGTTCCTGTTTCGGCCAAATTAGCTTCATCCAGGTCTTTAAAAGTCCTGAAGAATTC
CT
CCAGCAACGAGTGGACCGGTGAGACTGGCAAGGGAACCTTGCTGGACTCTGACGAGGGCAGCTTGAGCAGCGGCACCC
AG
AGCAGCGGCTGCCCCACGGCTGAGGGCAGGCAGGACTCCAAGAGCACTTCTCCCGGGAAATACCCACACCCCGGCTTG
GC
AGATTTTGCTGACAATCTCATCAAAGAGAGTGACATCCTGAGCGATGAAGATGATGACCACCGTCAGACTGTGAAGCA
GG
GCAGCCCTACTAAAGACATCGAAATTCAGTTCCAGAGACTGAGGATTTCCGAGGACCCAGACGTTCACCCCGAGGCTG
AG
CAGCAGCCTGGCCCGGAGTCGGGTGAGGGTCAGAAAGGAGGAGAGCAGCCCAAACTGGTCCGGGGGCACTTCTGCCCC
AT
TAAACGAAAAGCCAACAGCACCAAGAGGGACAGAGGAACTTTGCTCAAGGCGCAGATCCGTCACCAGTCCCTTGACAG
TC
AGTCTGAAAATGCCACCATCGACCTAAATTCTGTTCTAGAGCGAGAATTCAGTGTCCAGAGTTTAACATCTGTTGTCA
GT
GAGGAGTGTTTTTATGAAACAGAGAGCCACGGAAAATCATAGTATGATTCAATCCAGATATGGGTTAAATTCCTCATT
TT
ACTTTTAAACTGGTGGTAAAGTGGAAATTGCAAAAAAAAAAAAAAAAAAAACTGTTCATTCCTGGGTTTTGTGCAGTAT
AC
ATTTTCCCACAAAATGGTTGTAAAGATTTAAGTTATTTTAATTTATTGTGGATCAGAAACCTAGATGAAACTGGTCAG
AA
TCTGTAAATTACTTAGTTTATATCCACTTTGAGCAGGTATCAAATGATTTAGGATCCTTAAAATTACATTCTAATAAT
TA
AGTTATGTGGAAAAAGTAAGGCTGGGAAGTCGTGATTAATAGTTTTCAAAGGCCATTTTTTAAAATCCTCTGGGCATT
TT
CTTTCAGCTGTTTGTTAGTTTTTGCTTTATTTAAAGCATATTTAAGTTATTTTAATGTGGTTTAGGGGCAAAATGTGC
AG
ATACTTCATTTTTGTAAGATAGATTGTAATAGATGCTGTTTATACTAAACATGTCATAACTATCTATACAGTATATAT
TA
AAAGAAAGCTTGTACTGTATCTTATTTGATGATATTTATTTTCTCTGCCAAGCTGTATAGTAAAAGGAAAATAAGTCA
CA
TCTGGTCATTGGCATTTGTATCGTCATTCTGTAAAGACAAAAGAGTACCTATATAAGAAGCTCCACGTAGTGCAAATC

GA
CATCTGGTAGGCTGCTCGCCCCCAGGCAGCAGCTAGAGTCTGTAATTCTCTGCGTCATCCTCTTCTTTTTCTTCATTT
TT
GCTTTTTCTTCGCTTGAGTTCTTCTCTGAAATTATATGCAAAGAGTTGTGGGTCTTCATCACACATTTTTCTGTATAC
AT
CACAGAGGCTCTTAAAGTGTGAGATGGAGAGCTGGTGGGGCCGAAGAGTAGGGTCTATGTCTGCCAACTCTAACAGCC
TG
CCCGTGCTTTCCAAGCGCTGCGCTTCAGGGAATAACATTCTGAGCCCTCGATGGCAGTATTTCCTTCGGAACTGAAAT
AC
ATTCTGAACCACTTTTTCCACCAGCTTGAATGGCTGCTCTATCTTGGGCTGTATCAAGGGAGTGAAGTGCACCACGCC
CA
CGTCCACCTTCGTTGTAAGCAAACATATTATCATTCTGTGGCATGATATGTGGCATAGTGTGATCAATCAACTCATCC
TT
GTAAAACAGGAAGATGGGCTGTCAACAGCCTGTTTTCATAAACAGACCTTTCCACGTACTTCGGTTTCATCTCTAGGC
AT GGAAGATGGTACATTCTGGATTCGCAAATGACATGGAGAAATCAGCCGGCTGCACCTGTTCTCT

SEQ ID NO:9
4.4 with mini exon
exonTGCCGCGGCTGCTGAAGAACCAGACCAAAAACCTGCTTCAGAAGATAGACATGGACAGCAAGATGAAGAAGATG
GC
AGAGCTGCAGCTGTCCGTGGTGAGCGACCCAAAGAACAGGAAAGCCATAGAGAACCAGATCCAGCAATGGGAGCAGAA
TC
TTGAGAAATTTCACATGGATCTGTTCAGGATGCGCTGCTATCTGGCCAGCCTACAAGGTGGGGAGTTACCGAACCCAA
AG
AGTCTCCTTGCAGCCGCCAGCCGCCCCTCCAAGCCGGCCCTCGGCAGGCTGGGCATCTTGTCTGTTTCCTCTTTCCAT
GC
TCTGGTATGTTCTAGAGATGACTCTGCTCTCCGGAAAAGGACACTGTCACTGACCCAGCGAGGGAGAAACAAGAAGGG
AA
TATTTTCTTCGTTAAAAGGGCTGGACACACTGGCCAGAAAAGGCAAGGAGAAGAGACCTTCTATAACTCAGATATTTG
AT
TCAAGTGGCAGCCATGGATTTTCTGGAACTCAGCTACCTCAAAACTCCAGTAACTCCAGTGAGGTCGATGAACTTCTG
CA
TATATATGGTTCAACAGTAGACGGTGTTCCCCGAGACAATACATGGGAAATCCAGACTTATGTCCACTTTCAGGACAA
TC
ACGGAGTTACTGTAGGGATCAAGCCAGAGCACAGAGTAGAAGATATTTTGACTTTGGCATGCAAGATGAGGCAGTTGG
AA
CCCAGCCATTATGGCCTACAGCTTCGAAAATTAGTAGATGACAATGTTGAGTATTGCATCCCTGCACCATATGAATAT
AT
GCAACAACAGGTTTATGATGAAATAGAAGTCTTTCCACTAAATGTTTATGACGTGCAGCTCACGAAGACTGGGAGTGT

GT
GTGACTTTGGGTTTGCAGTTACAGCGCAGGTGGATGAGCGTCAGCATCTCAGCCGGATATTTATAAGCGACGTTCTTC
CC
GATGGCCTGGCGTATGGGGAAGGGCTGAGAAAGGGCAATGAGATCATGACCTTAAATGGGGAAGCTGTGTCTGATCTT
GA
CCTTAAGCAGATGGAGGCCCTGTTTTCTGAGAAGAGCGTCGGACTCACTCTGATTGCCCGGCCTCCGGACACAAAAGC
AA
CCCTGTGTACATCCTGGTCAGACAGTGACCTGTTCTCCAGGGACCAGAAGAGTCTGCTGCCCCCTCCTAACCAGTCCC
AA
CTGCTGGAGGAATTCCTGGATAACTTTAAAAAGAATACAGCCAATGATTTCAGCAACGTCCCTGATATCACAACAGGT
CT
GAAAAGGAGTCAGACAGATGGCACTCTGGATCAGGTTTCCCACAGGGAGAAAATGGAGCAGACATTCAGGAGTGCTGA
GC
AGATCACTGCACTGTGCAGGAGTTTTAACGACAGTCAGGCCAACGGCATGGAAGGACCGCGGGAGAATCAGGATCCTC
CT
CCGAGGCCTCTGGCCCGCCACCTGTCTGATGCAGACCGCCTCCGCAAAGTCATCCAGGAGCTTGTGGACACAGAGAAG
TC
CTACGTGAAGGATTTGAGCTGCCTCTTTGAATTATACTTGGAGCCACTTCAGAATGAGACCTTTCTTACCCAAGATGA
GA
TGGAGTCACTTTTTGGAAGTTTGCCAGAGATGCTTGAGTTTCAGAAGGTGTTTCTGGAGACCCTGGAGGATGGGATTT
CA
GCATCATCTGACTTTAACACCCTAGAAACCCCCTCACAGTTTAGAAAATTACTGTTTTCCCTTGGAGGCTCTTTCCTT
TA
TTACGCGGACCACTTTAAACTGTACAGTGGATTCTGTGCTAACCATATCAAAGTACAGAAGGTTCTGGAGCGAGCTAA
AA
CTGACAAAGCCTTCAAGGCTTTTCTGGACGCCCGGAACCCCACCAAGCAGCATTCCTCCACGCTGGAGTCCTACCTCA
TC
AAGCCGGTTCAGAGAGTGCTCAAGTACCCGCTGCTGCTCAAGGAGCTGGTGTCCCTGACGGACCAGGAGAGCGAGGAG
CA
CTACCACCTGACGGAAGCACTAAAGGCAATGGAGAAAGTAGCGAGCCACATCAATGAGATGCAGAAGATCTATGAGGA
TT
ATGGGACCGTGTTTGACCAGCTAGTAGCTGAGCAGAGCGGAACAGAGAAGGAGGTAACAGAACTTTCGATGGGAGAGC
TT
CTGATGCACTCTACGGTTTCCTGGTTGAATCCATTTCTGTCTCTAGGAAAAGCTAGAAAGGACCTTGAGCTCACAGTA
TT
TGTTTTTAAGAGAGCCGTCATACTGGTTTATAAAGAAAACTGCAAACTGAAAAAGAAATTGCCCTCGAATTCCCGGCC
TG
CACACAACTCTACTGACTTGGACCCATTTAAATTCCGCTGGTTGATCCCCATCTCCGCGCTTCAAGTCAGACTGGGGA
AT
CCAGCAGGGACAGAAAATAATTCCATATGGGAACTGATCCATACGAAGTCAGAAATAGAAGGACGGCCAGAAACCATC

TT
TCAGTTGTGTTGCAGTGACAGTGAAAGCAAAACCAACATTGTTAAGGTGATTCGTTCTATTCTGAGGGAGAACTTCAG
GC
GTCACATAAAGTGTGAATTACCACTGGAGAAAACGTGTAAGGATCGCCTGGTACCTCTTAAGAACCGAGTTCCTGTTT
CG
GCCAAATTAGCTTCATCCAGGTCTTTAAAAGTCCTGAAGAATTCCTCCAGCAACGAGTGGACCGGTGAGACTGGCAAG
GG
AACCTTGCTGGACTCTGACGAGGGCAGCTTGAGCAGCGGCACCCAGAGCAGCGGCTGCCCCACGGCTGAGGGCAGGCA
GG
ACTCCAAGAGCACTTCTCCCGGGAAATACCCACACCCCGGCTTGGCAGATTTTGCTGACAATCTCATCAAAGAGAGTG
AC
ATCCTGAGCGATGAAGATGATGACCACCGTCAGACTGTGAAGCAGGGCAGCCCTACTAAAGACATCGAAATTCAGTTC
CA
GAGACTGAGGATTTCCGAGGACCCAGACGTTCACCCCGAGGCTGAGCAGCAGCCTGGCCCGGAGTCGGGTGAGGGTCA
GA
AAGGAGGAGAGCAGCCCAAACTGGTCCGGGGGCACTTCTGCCCCATTAAACGAAAAGCCAACAGCACCAAGAGGGACA
GA
GGAACTTTGCTCAAGGCGCAGATCCGTCACCAGTCCCTTGACAGTCAGTCTGAAAATGCCACCATCGACCTAAATTCT
GT
TCTAGAGCGAGAATTCAGTGTCCAGAGTTTAACATCTGTTGTCAGTGAGGAGTGTTTTTATGAAACAGAGAGCCACGG
AA
AATCATAGTATGATTCAATCCAGATATGGGTTAAATTCCTCATTTTACTTTTAAACTGGTGGTAAAGTGGAAATTGCA
AA
AAAAAAAAAAAAAAAACTGTTCATTCCTGGGTTTTGTGCAGTATACATTTTCCCACAAAATGGTTGTAAAGATTTAAG
TT
ATTTTAATTTATTGTGGATCAGAAACCTAGATGAAACTGGTCAGAATCTGTAAATTACTTAGTTTATATCCACTTTGA
GC
AGGTATCAAATGATTTAGGATCCTTAAAATTACATTCTAATAATTAAGTTATGTGGAAAAAGTAAGGCTGGGAAGTCG
TG
ATTAATAGTTTTCAAAGGCCATTTTTTAAAATCCTCTGGGCATTTTCTTTCAGCTGTTTGTTAGTTTTTGCTTTATTT
AA
AGCATATTTAAGTTATTTTAATGTGGTTTAGGGGCAAAATGTGCAGATACTTCATTTTTGTAAGATAGATTGTAATAG
AT
GCTGTTTATACTAAACATGTCATAACTATCTATACAGTATATATTAAAAGAAAGCTTGTACTGTATCTTATTTGATGA
TA
TTTATTTTCTCTGCCAAGCTGTATAGTAAAAGGAAAATAAGTCACATCTGGTCATTGGCATTTGTATCGTCATTCTGT
AA
AGACAAAAGAGTACCTATATAAGAAGCTCCACGTAGTGCAAATCGACATCTGGTAGGCTGCTCGCCCCCAGGCAGCAG
CT
AGAGTCTGTAATTCTCTGCGTCATCCTCTTCTTTTTCTTCATTTTTGCTTTTTCTTCGCTTGAGTTCTTCTCTGAAAT

TA
TATGCAAAGAGTTGTGGGTCTTCATCACACATTTTTCTGTATACATCACAGAGGCTCTTAAAGTGTGAGATGGAGAGC
TG
GTGGGGCCGAAGAGTAGGGTCTATGTCTGCCAACTCTAACAGCCTGCCCGTGCTTTCCAAGCGCTGCGCTTCAGGGAA
TA
ACATTCTGAGCCCTCGATGGCAGTATTTCCTTCGGAACTGAAATACATTCTGAACCACTTTTTCCACCAGCTTGAATG
GC
TGCTCTATCTTGGGCTGTATCAAGGGAGTGAAGTGCACCACGCCCACGTCCACCTTCGTTGTAAGCAAACATATTATC
AT
TCTGTGGCATGATATGTGGCATAGTGTGATCAATCAACTCATCCTTGTAAAACAGGAAGATGGGCTGTCAACAGCCTG
TT
TTCATAAACAGACCTTTCCACGTACTTCGGTTTCATCTCTAGGCATGGAAGATGGTACATTCTGGATTCGCAAATGAC
AT GGAGAAATCAGCCGGCTGCACCTGTTCTCT

SEQ ID NO:10

4.4 with out mini exon

TGCCGCGGCTGCTGAAGAACCAGACCAAAAACCTGCTTCAGAAGATAGACATGGACAGCAAGATGAAGAAGATGGCAG
AG
CTGCAGCTGTCCGTGGTGAGCGACCCAAAGAACAGGAAAGCCATAGAGAACCAGATCCAGCAATGGGAGCAGAATCTT
GA
GAAATTTCACATGGATCTGTTCAGGATGCGCTGCTATCTGGCCAGCCTACAAGGTGGGGAGTTACCGAACCCAAAGAG
TC
TCCTTGCAGCCGCCAGCCGCCCCTCCAAGCCGGCCCTCGGCAGGCTGGGCATCTTGTCTGTTTCCTCTTTCCATGCTC
TG
GTATGTTCTAGAGATGACTCTGCTCTCCGGAAAAGGACACTGTCACTGACCCAGCGAGGGAGAAACAAGAAGGGAATA
TT
TTCTTCGTTAAAAGGGCTGGACACACTGGCCAGAAAAGGCAAGGAGAAGAGACCTTCTATAACTCAGGTCGATGAACT
TC
TGCATATATATGGTTCAACAGTAGACGGTGTTCCCCGAGACAATACATGGGAAATCCAGACTTATGTCCACTTTCAGG
AC
AATCACGGAGTTACTGTAGGGATCAAGCCAGAGCACAGAGTAGAAGATATTTTGACTTTGGCATGCAAGATGAGGCAG
TT
GGAACCCAGCCATTATGGCCTACAGCTTCGAAAATTAGTAGATGACAATGTTGAGTATTGCATCCCTGCACCATATGA
AT
ATATGCAACAACAGGTTTATGATGAAATAGAAGTCTTTCCACTAAATGTTTATGACGTGCAGCTCACGAAGACTGGGA
GT
GTGTGTGACTTTGGGTTTGCAGTTACAGCGCAGGTGGATGAGCGTCAGCATCTCAGCCGGATATTTATAAGCGACGTT
CT
TCCCGATGGCCTGGCGTATGGGGAAGGGCTGAGAAAGGGCAATGAGATCATGACCTTAAATGGGGAAGCTGTGTCTGA

TC
TTGACCTTAAGCAGATGGAGGCCCTGTTTTCTGAGAAGAGCGTCGGACTCACTCTGATTGCCCGGCCTCCGGACACAA
AA
GCAACCCTGTGTACATCCTGGTCAGACAGTGACCTGTTCTCCAGGGACCAGAAGAGTCTGCTGCCCCCTCCTAACCAG
TC
CCAACTGCTGGAGGAATTCCTGGATAACTTTAAAAAGAATACAGCCAATGATTTCAGCAACGTCCCTGATATCACAAC
AG
GTCTGAAAAGGAGTCAGACAGATGGCACTCTGGATCAGGTTTCCCACAGGGAGAAAATGGAGCAGACATTCAGGAGTG
CT
GAGCAGATCACTGCACTGTGCAGGAGTTTTAACGACAGTCAGGCCAACGGCATGGAAGGACCGCGGGAGAATCAGGAT
CC
TCCTCCGAGGCCTCTGGCCCGCCACCTGTCTGATGCAGACCGCCTCCGCAAAGTCATCCAGGAGCTTGTGGACACAGA
GA
AGTCCTACGTGAAGGATTTGAGCTGCCTCTTTGAATTATACTTGGAGCCACTTCAGAATGAGACCTTTCTTACCCAAG
AT
GAGATGGAGTCACTTTTTGGAAGTTTGCCAGAGATGCTTGAGTTTCAGAAGGTGTTTCTGGAGACCCTGGAGGATGGG
AT
TTCAGCATCATCTGACTTTAACACCCTAGAAACCCCCTCACAGTTTAGAAAATTACTGTTTTCCCTTGGAGGCTCTTT
CC
TTTATTACGCGGACCACTTTAAACTGTACAGTGGATTCTGTGCTAACCATATCAAAGTACAGAAGGTTCTGGAGCGAG
CT
AAAACTGACAAAGCCTTCAAGGCTTTTCTGGACGCCCGGAACCCCACCAAGCAGCATTCCTCCACGCTGGAGTCCTAC
CT
CATCAAGCCGGTTCAGAGAGTGCTCAAGTACCCGCTGCTGCTCAAGGAGCTGGTGTCCCTGACGGACCAGGAGAGCGA
GG
AGCACTACCACCTGACGGAAGCACTAAAGGCAATGGAGAAAGTAGCGAGCCACATCAATGAGATGCAGAAGATCTATG
AG
GATTATGGGACCGTGTTTGACCAGCTAGTAGCTGAGCAGAGCGGAACAGAGAAGGAGGTAACAGAACTTTCGATGGGA
GA
GCTTCTGATGCACTCTACGGTTTCCTGGTTGAATCCATTTCTGTCTCTAGGAAAAGCTAGAAAGGACCTTGAGCTCAC
AG
TATTTGTTTTTAAGAGAGCCGTCATACTGGTTTATAAAGAAAACTGCAAACTGAAAAAGAAATTGCCCTCGAATTCCC
GG
CCTGCACACAACTCTACTGACTTGGACCCATTTAAATTCCGCTGGTTGATCCCCATCTCCGCGCTTCAAGTCAGACTG
GG
GAATCCAGCAGGGACAGAAAATAATTCCATATGGGAACTGATCCATACGAAGTCAGAAATAGAAGGACGGCCAGAAAC
CA
TCTTTCAGTTGTGTTGCAGTGACAGTGAAAGCAAAACCAACATTGTTAAGGTGATTCGTTCTATTCTGAGGGAGAACT
TC
AGGCGTCACATAAAGTGTGAATTACCACTGGAGAAAACGTGTAAGGATCGCCTGGTACCTCTTAAGAACCGAGTTCCT

GT
TTCGGCCAAATTAGCTTCATCCAGGTCTTTAAAAGTCCTGAAGAATTCCTCCAGCAACGAGTGGACCGGTGAGACTGG
CA
AGGGAACCTTGCTGGACTCTGACGAGGGCAGCTTGAGCAGCGGCACCCAGAGCAGCGGCTGCCCCACGGCTGAGGGCA
GG
CAGGACTCCAAGAGCACTTCTCCCGGGAAATACCCACACCCCGGCTTGGCAGATTTTGCTGACAATCTCATCAAAGAG
AG
TGACATCCTGAGCGATGAAGATGATGACCACCGTCAGACTGTGAAGCAGGGCAGCCCTACTAAAGACATCGAAATTCA
GT
TCCAGAGACTGAGGATTTCCGAGGACCCAGACGTTCACCCCGAGGCTGAGCAGCAGCCTGGCCCGGAGTCGGGTGAGG
GT
CAGAAAGGAGGAGAGCAGCCCAAACTGGTCCGGGGGCACTTCTGCCCCATTAAACGAAAAGCCAACAGCACCAAGAGG
GA
CAGAGGAACTTTGCTCAAGGCGCAGATCCGTCACCAGTCCCTTGACAGTCAGTCTGAAAATGCCACCATCGACCTAAA
TT
CTGTTCTAGAGCGAGAATTCAGTGTCCAGAGTTTAACATCTGTTGTCAGTGAGGAGTGTTTTTATGAAACAGAGAGCC
AC
GGAAAATCATAGTATGATTCAATCCAGATATGGGTTAAATTCCTCATTTTACTTTTAAACTGGTGGTAAAGTGGAAAT
TG
CAAAAAAAAAAAAAAAAAAACTGTTCATTCCTGGGTTTTGTGCAGTATACATTTTCCCACAAAATGGTTGTAAAGATT
TA
AGTTATTTTAATTTATTGTGGATCAGAAACCTAGATGAAACTGGTCAGAATCTGTAAATTACTTAGTTTATATCCACT
TT
GAGCAGGTATCAAATGATTTAGGATCCTTAAAATTACATTCTAATAATTAAGTTATGTGGAAAAAGTAAGGCTGGGAA
GT
CGTGATTAATAGTTTTCAAAGGCCATTTTTTAAAATCCTCTGGGCATTTTCTTTCAGCTGTTTGTTAGTTTTTGCTTT
AT
TTAAAGCATATTTAAGTTATTTTAATGTGGTTTAGGGGCAAAATGTGCAGATACTTCATTTTTGTAAGATAGATTGTA
AT
AGATGCTGTTTATACTAAACATGTCATAACTATCTATACAGTATATATTAAAAGAAAGCTTGTACTGTATCTTATTTG
AT
GATATTTATTTTCTCTGCCAAGCTGTATAGTAAAAGGAAAATAAGTCACATCTGGTCATTGGCATTTGTATCGTCATT
CT
GTAAAGACAAAAGAGTACCTATATAAGAAGCTCCACGTAGTGCAAATCGACATCTGGTAGGCTGCTCGCCCCCAGGCA
GC
AGCTAGAGTCTGTAATTCTCTGCGTCATCCTCTTCTTTTTCTTCATTTTTGCTTTTTCTTCGCTTGAGTTCTTCTCTG
AA
ATTATATGCAAAGAGTTGTGGGTCTTCATCACACATTTTTCTGTATACATCACAGAGGCTCTTAAAGTGTGAGATGGA
GA
GCTGGTGGGGCCGAAGAGTAGGGTCTATGTCTGCCAACTCTAACAGCCTGCCCGTGCTTTCCAAGCGCTGCGCTTCAG

GG
AATAACATTCTGAGCCCTCGATGGCAGTATTTCCTTCGGAACTGAAATACATTCTGAACCACTTTTTCCACCAGCTTG
AA
TGGCTGCTCTATCTTGGGCTGTATCAAGGGAGTGAAGTGCACCACGCCCACGTCCACCTTCGTTGTAAGCAAACATAT
TA
TCATTCTGTGGCATGATATGTGGCATAGTGTGATCAATCAACTCATCCTTGTAAAACAGGAAGATGGGCTGTCAACAG
CC
TGTTTTCATAAACAGACCTTTCCACGTACTTCGGTTTCATCTCTAGGCATGGAAGATGGTACATTCTGGATTCGCAAA
TG ACATGGAGAAATCAGCCGGCTGCACCTGTTCTCT

3.3 amino acid sequence = SEQ ID NO:11

MEGPRENQDPPPRPLARHLSDADRLRKVIQELVDTEKSYVKDLSCLFELYLEPLQNETFLTQDEMESLFGSLPEMLEF
QK
VFLETLEDGISASSDFNTLETPSQFRKLLFSLGGSFLYYADHFKLYSGFCANHIKVQKVLERAKTDKAFKAFLDARNP
TK
QHSSTLESYLIKPVQRVLKYPLLLKELVSLTDQESEEHYHLTEALKAMEKVASHINEMQKIYEDYGTVFDQLVAEQSG
TE
KEVTELSMGELLMHSTVSWLNPFLSLGKARKDLELTVFVFKRAVILVYKENCKLKKKLPSNSRPAHNSTDLDPFKFRW
LI
PISALQVRLGNPAGTENNSIWELIHTKSEIEGRPETIFQLCCSDSESKTNIVKVIRSILRENFRRHIKCELPLEKTCK
DR
LVPLKNRVPVSAKLASSRSLKVLKNSSSNEWTGETGKGTLLDSDEGSLSSGTQSSGCPTAEGRQDSKSTSPGKYPHPG
LA
DFADNLIKESDILSDEDDDHRQTVKQGSPTKDIEIQFQRLRISEDPDVHPEAEQQPGPESGEGQKGGEQPKLVRGHFC
PI KRKANSTKRDRGTLLKAQIRHQSLDSQSENATIDLNSVLEREFSVQSLTSVVSEECFYETESHGKS

4.4 without exon amino acid sequence = SEQ ID NO:12

MDSKMKKMAELQLSVVSDPKNRKAIENQIQQWEQNLEKFHMDLFRMRCYLASLQGGELPNPKSLLAAASRPSKPALGR
LG
ILSVSSFHALVCSRDDSALRKRTLSLTQRGRNKKGIFSSLKGLDTLARKGKEKRPSITQVDELLHIYGSTVDGVPRDN
TW
EIQTYVHFQDNHGVTVGIKPEHRVEDILTLACKMRQLEPSHYGLQLRKLVDDNVEYCIPAPYEYMQQQVYDEIEVFPL
NV
YDVQLTKTGSVCDFGFAVTAQVDERQHLSRIFISDVLPDGLAYGEGLRKGNEIMTLNGEAVSDLDLKQMEALFSEKSV
GL
TLIARPPDTKATLCTSWSDSDLFSRDQKSLLPPPNQSQLLEEFLDNFKKNTANDFSNVPDITTGLKRSQTDGTLDQVS

HR
EKMEQTFRSAEQITALCRSFNDSQANGMEGPRENQDPPPRPLARHLSDADRLRKVIQELVDTEKSYVKDLSCLFELYL
EP
LQNETFLTQDEMESLFGSLPEMLEFQKVFLETLEDGISASSDFNTLETPSQFRKLLFSLGGSFLYYADHFKLYSGFCA
NH
IKVQKVLERAKTDKAFKAFLDARNPTKQHSSTLESYLIKPVQRVLKYPLLLKELVSLTDQESEEHYHLTEALKAMEKV
AS
HINEMQKIYEDYGTVFDQLVAEQSGTEKEVTELSMGELLMHSTVSWLNPFLSLGKARKDLELTVFVFKRAVILVYKEN
CK
LKKKLPSNSRPAHNSTDLDPFKFRWLIPISALQVRLGNPAGTENNSIWELIHTKSEIEGRPETIFQLCCSDSESKTNI
VK
VIRSILRENFRRHIKCELPLEKTCKDRLVPLKNRVPVSAKLASSRSLKVLKNSSSNEWTGETGKGTLLDSDEGSLSSG
TQ
SSGCPTAEGRQDSKSTSPGKYPHPGLADFADNLIKESDILSDEDDDHRQTVKQGSPTKDIEIQFQRLRISEDPDVHPE
AE
QQPGPESGEGQKGGEQPKLVRGHFCPIKRKANSTKRDRGTLLKAQIRHQSLDSQSENATIDLNSVLEREFSVQSLTSV
VS EECFYETESHGKS

4.4 with exon amino acid sequence = SEQ ID NO:13

MDSKMKKMAELQLSVVSDPKNRKAIENQIQQWEQNLEKFHMDLFRMRCYLASLQGGELPNPKSLLAAASRPSKPALGR
LG
ILSVSSFHALVCSRDDSALRKRTLSLTQRGRNKKGIFSSLKGLDTLARKGKEKRPSITQIFDSSGSHGFSGTQLPQNS
SN
SSEVDELLHIYGSTVDGVPRDNTWEIQTYVHFQDNHGVTVGIKPEHRVEDILTLACKMRQLEPSHYGLQLRKLVDDNV
EY
CIPAPYEYMQQQVYDEIEVFPLNVYDVQLTKTGSVCDFGFAVTAQVDERQHLSRIFISDVLPDGLAYGEGLRKGNEIM
TL
NGEAVSDLDLKQMEALFSEKSVGLTLIARPPDTKATLCTSWSDSDLFSRDQKSLLPPPNQSQLLEEFLDNFKKNTAND
FS
NVPDITTGLKRSQTDGTLDQVSHREKMEQTFRSAEQITALCRSFNDSQANGMEGPRENQDPPPRPLARHLSDADRLRK
VI
QELVDTEKSYVKDLSCLFELYLEPLQNETFLTQDEMESLFGSLPEMLEFQKVFLETLEDGISASSDFNTLETPSQFRK
LL
FSLGGSFLYYADHFKLYSGFCANHIKVQKVLERAKTDKAFKAFLDARNPTKQHSSTLESYLIKPVQRVLKYPLLLKEL
VS
LTDQESEEHYHLTEALKAMEKVASHINEMQKIYEDYGTVFDQLVAEQSGTEKEVTELSMGELLMHSTVSWLNPFLSLG
KA
RKDLELTVFVFKRAVILVYKENCKLKKKLPSNSRPAHNSTDLDPFKFRWLIPISALQVRLGNPAGTENNSIWELIHTK
SE

IEGRPETIFQLCCSDSESKTNIVKVIRSILRENFRRHIKCELPLEKTCKDRLVPLKNRVPVSAKLASSRSLKVLKNSS
SN
EWTGETGKGTLLDSDEGSLSSGTQSSGCPTAEGRQDSKSTSPGKYPHPGLADFADNLIKESDILSDEDDDHRQTVKQG
SP
TKDIEIQFQRLRISEDPDVHPEAEQQPGPESGEGQKGGEQPKLVRGHFCPIKRKANSTKRDRGTLLKAQIRHQSLDSQ
SE NATIDLNSVLEREFSVQSLTSVVSEECFYETESHGKS