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

(51) International Patent Classification⁷: **C07K**

(21) International Application Number: PCT/US02/31618

(22) International Filing Date: 2 October 2002 (02.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/327,212   3 October 2001 (03.10.2001)   US

(71) Applicant *(for all designated States except US)*: **RIGEL PHARMACEUTICALS, INC.** [US/US]; 240 East Grand Avenue, South San Francisco, CA 94080 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **CHU, Peter** [US/US]; 15 Lincoln Way #105, San Francisco, CA 94122 (US). **LI, Congfen** [US/US]; 2115 Everglades Place, Davis, CA 95616 (US). **LIAO, X., Charlene** [US/US];

3751 Middlefield Road, Palo Alto, CA 94303 (US). **MASUDA, Esteban** [CO/US]; 232 Oakhurst Place, Menlo Park, CA 94025 (US). **PARDO, Jorge** [US/US]; 1379 MASONIC AVE., San Francisco, CA 94116 (US). **ZHAO, Haoran** [US/US]; 330 Port Royal Avenue, Foster City, CA 94404 (US).

(74) Agents: PARENT, Annette, S. et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

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

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

(54) Title: MODULATORS OF LYMPHOCYTE ACTIVATION AND MIGRATION

Identification of Regulatory Proteins that Affect T-Cell Activation
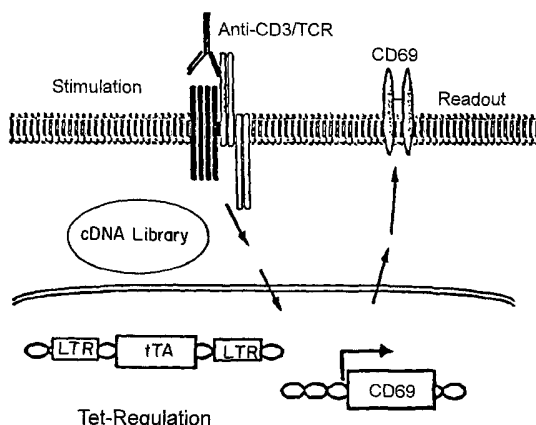
Rationale

CD69 is an endogenous gene whose expression is upregulated by stimulation through the T Cell Receptor (TCR) complex. After introducing a cDNA library to the T cells through retroviral infection, we asked whether any library members would inhibit the upregulation of CD69. Library members capable of blocking CD69 induction may achieve this through inhibition of the TCR signal transduction pathway.

(57) Abstract: The present invention relates to regulation of lymphocyte activation and migration. More particularly, the present invention is directed to nucleic acids encoding the nucleic acids and proteins listed in Figure 7, which are involved in modulation of lymphocyte activation and migration, e.g., A-raf-1, Lck, Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH. The invention further relates to methods for identifying and using agents, including small organic molecules, antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, siRNA, and ribozymes, that modulate lymphocyte activation or migration; as well as to the use of expression profiles and compositions in diagnosis and therapy related to lymphocyte activation and suppression, and lymphocyte migration.

ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

MODULATORS OF LYMPHOCYTE ACTIVATION AND MIGRATION

CROSS-REFERENCES TO RELATED APPLICATIONS

The present application claims priority to USSN 60/327,212, filed October 3, 2001, herein incorporated by reference in its entirety.

The present application is related to PCT/US02/12342; USSN 09/971,28, filed October 3, 2001; PCT/US02/11205; USSN 09/998,667, filed November 30, 2001; PCT/US02/10257; USSN 09/967,624, filed September 28, 2001; PCT/US/17417; USSN 10/160,354, filed May 30, 2002; USSN 60/362,034, filed March 4, 2002; and USSN 10/233,098, filed August 30, 2002, herein each incorporated by reference in their entirety.

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

Not applicable.

FIELD OF THE INVENTION

The present invention relates to regulation of lymphocyte activation and migration. More particularly, the present invention is directed to nucleic acids encoding the nucleic acids and proteins listed in Figure 7, which are involved in modulation of lymphocyte activation and migration, e.g., A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH. The invention further relates to methods for identifying and using agents, including small organic molecules, antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, siRNA, and ribozymes, that modulate lymphocyte activation or migration; as well as to the use of expression profiles and compositions in diagnosis and therapy related to lymphocyte activation and suppression, and lymphocyte migration.

BACKGROUND OF THE INVENTION

The immune response includes both a cellular and a humoral response. The cellular response is mediate largely by T lymphocytes (alternatively and equivalently referred to herein as T cells), while the humoral response is mediated by B lymphocytes (alternatively and equivalently referred to herein as B cells). Lymphocytes play a number of crucial roles in immune responses, including direct killing of virus-infected cells, cytokine and antibody production, and facilitation of B cell responses. Lymphocytes are also involved in acute and chronic inflammatory disease; asthma; allergies; autoimmune diseases such as scleroderma, pernicious anemia, multiple sclerosis, myasthenia gravis, IDDM, rheumatoid arthritis, systemic lupus erythematosus, and Crohn's disease; and organ and tissue transplant disease, e.g., graft vs. host disease.

B lymphocytes produce and secrete antibodies in response to the concerted presentation of antigen and MHC class II molecules on the surface of antigen presenting cells. Antigen presentation initiates B cell activation through the B cell receptor (BCR) at the B cell surface. Signal transduction from the BCR leads to B cell activation and changes in B cell gene expression, physiology, and function, including secretion of antibodies.

T cells do not produce antibodies, but many subtypes of T cells produce co-stimulatory molecules that augment antibody production by B cells during the humoral immune response. In addition, many T cells engulf and destroy cells or agents that are recognized by cell surface receptors. Engagement of the cell surface T cell receptor (TCR) initiates T cell activation. Signal transduction from the TCR leads to T cell activation and changes in T cell gene expression, physiology, and function, including the secretion of cytokines.

Identifying ligands, receptors, and signaling proteins downstream of TCR, as well as BCR, activation is important for developing therapeutic regents to inhibit immune response in inflammatory disease, autoimmune disease, and organ transplant, as well as to activate immune response in immunocompromised subjects, and in patients with infectious disease and cancer (*see, e.g.,* Rogge *et al., Nature Genetics* 25:96-101 (2000)). In addition, identification of molecules participating in lymphocyte migration is important for developing therapeutic reagents, as described above,

SUMMARY OF THE INVENTION

The present invention therefore provides nucleic acids and proteins, as shown in Figure 7 and the sequence listing provided herein, which are involved in modulation of

lymphocyte activation and migration, e.g., A-raf-1, Lck, Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228,

5   Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH. The invention therefore provides methods of screening for compounds, e.g., small organic molecules, antibodies, peptides, lipids, peptides, cyclic peptides, nucleic acids, antisense molecules, siRNA, and ribozyme, that are capable of modulating lymphocyte activation and lymphocyte migration, e.g., either activating or inhibiting lymphocytes and their ability to

10   migrate. Therapeutic and diagnostic methods and reagents are also provided.

In one aspect of the invention, nucleic acids encoding A-raf-1, Lck, Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70,

15   SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH are provided. In another aspect, the present invention provides nucleic acids, such as probes, antisense oligonucleotides, siRNA, and ribozymes, that hybridize to a gene encoding a protein as listed in Figure 7. In another aspect, the invention provides expression vectors and host cells comprising nucleic acids encoding proteins listed

20   in Figure 7. In another aspect, the present invention provides the proteins listed in Figure 7, and antibodies thereto.

In another aspect, the present invention provides a method for identifying a compound that modulates lymphocyte activation or lymphocyte migration, the method comprising the steps of: (i) contacting a cell comprising an A-raf-1, Lck, Zap70, Syk, PLCγl,

25   PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, or RNH polypeptide or fragment thereof with the compound, the polypeptide or fragment

30   thereof encoded by a nucleic acid that hybridizes under stringent conditions to a nucleic acid comprising a nucleotide sequence of A-raf-1, Lck, Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from

clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, or RNH; and (ii)
determining the chemical or phenotypic effect of the compound upon the cell comprising the
polypeptide or fragment thereof, thereby identifying a compound that modulates lymphocyte
activation or migration.

5   In another aspect, the present invention provides a method for identifying a
compound that modulates lymphocyte activation or migration, the method comprising the
steps of: (i) contacting the compound with a A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG,
SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα,
integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7,
10   SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228,
Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, or RNH
polypeptide or a fragment thereof, the polypeptide or fragment thereof encoded by a nucleic
acid that hybridizes under stringent conditions to a nucleic acid comprising a nucleotide
sequence of A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP,
15   PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2,
P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S,
RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin,
filamin A α, centractin α, moesin, TIMP3, or RNH; (ii) determining the physical effect of the
compound upon the polypeptide; and (iii) determining the chemical or phenotypic effect of
20   the compound upon a cell comprising an polypeptide or fragment thereof, thereby identifying
a compound that modulates lymphocyte activation or migration.

          In one embodiment, the polypeptide or fragment thereof is encoded by a
nucleic acid that hybridizes under stringent conditions to a nucleic acid comprising a
sequence as listed in Table 7 or the sequence listing herein.

25   In another embodiment, the host cell is a B lymphocyte or a T lymphocyte. In
another embodiment, the host cell is a primary or cultured cell, e.g., a BJAB or Jurkat cell.

          In one embodiment, the chemical or phenotypic effect is determined by
measuring CD69 expression, IL-2 production, intracellular Ca2+ mobilization, or lymphocyte
proliferation.

30   In another embodiment, modulation is inhibition of T or B lymphocyte
activation or migration.

          In another embodiment, the polypeptide is recombinant.

          In another embodiment, the compound is an antibody, an antisense molecule,
an siRNA, a peptide, a circular peptide, or a small organic molecule.

4

In one embodiment, the chemical or phenotypic effect is determined by measuring lymphocyte migration *in vitro* toward a ligand, e.g., an EDG ligand such as SPP or LPA.

5   In one aspect, the present invention provides a method of modulating lymphocyte activation or migration in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified using the methods described above.

In one embodiment, the subject is a human.

In another aspect, the present invention provides a composition comprising a

10   therapeutically effective amount of an analog of 2-amino-2(2-[4-octylphenyl]ethyl)-1,3-propanediol hydrochloride and a physiologically acceptable carrier.

In one embodiment, the present invention provides method of modulating lymphocyte activation or migration in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of A-raf-1, Lck, Zap70, Syk,

15   PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, or RNH polypeptide, the polypeptide encoded by a nucleic acid that hybridizes under

20   stringent conditions to a nucleic acid comprising a nucleotide sequence of A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α,

25   centractin α, moesin, TIMP3, or RNH.

In another aspect, the present invention provides a method of modulating lymphocyte activation or migration in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a nucleic acid encoding A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1,

30   TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, or RNH or fragment thereof, wherein the nucleic acid hybridizes under stringent conditions to a nucleic acid encoding a polypeptide comprising a

5

nucleotide sequence of A-raf-1, Lck, Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK,

nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase la,

PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1,

TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone

5   2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, or RNH.

In another embodiment, the EDG nucleic acid is selected from the sequences

listed in Figure 7 or the sequence listing herein.

In one aspect, the present invention provides a method of modulating T

lymphocyte migration and activation in a subject, the method comprising the step of

10   administering to the subject a therapeutically effective amount of a first compound identified

using the methods described above, which first compound modulates activation, and

administering to the subject a therapeutically effective amount of a second compound

identified using the methods described above, which second compound modulates migration.

In another aspect, the present invention provides a method of modulating T

15   lymphocyte migration and activation in a subject, the method comprising the step of

administering to the subject a therapeutically effective amount of a compound identified

using the methods described above, which compounds modulates both activation and

migration.

20   BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic of identification of regulatory proteins that affect

T cell activation.

Figure 2 shows a schematic of TCR activation-induced expression of CD69.

Figure 3 shows a schematic of the distinction between cDNA-induced

25   phenotypes and somatic mutations.

Figure 4 shows a schematic of cell specificity of potential targets.

Figure 5 shows known TCR regulators identified from a CD69 cDNA screen.

Figure 6 shows primary, novel TCR regulators identified from a CD69 cDNA

screen.

30   Figure 7 provides a list of nucleic acids and the proteins that they encode.

These proteins were identified as modulators of lymphocyte activation and migration. These

proteins therefore can be used as molecular targets for discovery of compounds that modulate

lymphocyte activation and migration.

DETAILED DESCRIPTION OF THE INVENTION

Introduction

For the first time, the proteins listed in Figure 7 have been functionally identified as a protein involved in regulating lymphocyte activation and migration. These proteins were identified in a functional genetic screen using CD 69 as a readout of lymphocyte activation. These results indicate thatA-raf-1, Lck, Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH can be used for inhibition or activation of TCR and BCR signaling and lymphocyte activation. In one embodiment, modulators of these proteins are used to inhibit lymphocyte activation. In one embodiment, agonists of these proteins are used for inhibition of lymphocyte activation.

These results also indicate that EDG-1, other EDG family members such as EDG-5, and EDG modulators, e.g., antagonists or agonists, can be used for inhibition or activation of lymphocyte migration. In one embodiment, modulators of EDG family proteins are used to inhibit lymphocyte migration. In one embodiment, antagonists of EDG-1 are used for inhibition of lymphocyte migration.

Previously, EDG family proteins were known to be G-protein coupled receptors (GPCR, *see, e.g.*, WO 94/05695 and US Patent 5,508,384) that are expressed in a wide variety of cells (*see, e.g.*, Goetzl *et al., J. Immunol.* 164:4669-4999 (2000)). However, the function of EDG proteins was unknown. EDG-1 was identified as expressed in endothelial cells as well as in many other cells, and a role in angiogenesis has been proposed for this protein (*see, e.g.*, WO 91/15583; Bornfeldt *et al., J. Cell Biol.* 130:193-206 (1995); and Wang *et al., J. Biol. Chem.* 274:35343-35350 (1999)). It has also been speculated that EDG-1 is involved in numerous diverse disease states (*see, e.g.*, WO 99/46277). EDG-1 is ubiquitously expressed. EDG-4 has been identified as expressed in T lymphocytes, among other cells (*see, e.g.*, Goetzl *et al., J. Immunol.* 164:4669-4999 (2000)). A role for EDG-2 and other EDG family members in apoptosis, e.g., in lymphocytes, has also been proposed (*see, e.g.*, WO 99/19513).

EDG-1 and other EDG family members EDG-2 to -8 were known to bind sphingolipid ligands, e.g., sphingosine-1-phosphate (SPP, EDG-1, 3, 5, 6, and 8) or lysophosphatidic acid (LPA), EDG-2, 4, and 7) (*see, e.g.*, Okamoto *et al., J. Biol. Chem.* 273:27104-27110 (1998); Lee *et al., Science* 279:1552-1555 (1998); Lee *et al., J. Biol. Chem.*

273:22105-22112 (1998); Pyne & Pyne, *Biochem. J.* 349:385-402 (2000); and Windh *et al.*,
J. Biol. Chem. 274:27351-27358 (1999); and Prieschl & Baumruker, *Immunology Today*
21:555-560 (2000)). Recent screening for immunosuppressants has re-identified myriocin, a
sphingosine-like natural fungal product (Chen *et al.*, *Chemistry & Biology* 6:221-235 (1999)).
5   FTY720 is a synthetic analog of myriocin and has immunosuppressant activity, e.g., for
organ transplant and graft vs. host disease (2-amino-2(2-[4-octylphenyl]ethyl)-1,3-
propanediol hydrochloride). Its primary molecular target, however, is unknown (*see, e.g.*,
Brinkmann *et al.*, *TIPS* 21:49-52 (2000); Pinschewer *et al.*, *J. Immunol* 164:5761-5770
(2000)). Although extracellular ligands SPP and LPA were known to bind to EDG proteins,
10   the function of the EDG proteins remained unknown.

The present invention, therefore, has functionally identified A-raf-1, Lck,
Zap70, Syk, PLCγl, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1,
TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL,
DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD,
15   Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α,
centractin α, moesin, TIMP3, and RNH as drug targets for compounds that suppress or
activate lymphocyte activation and migration, e.g., for the treatment of diseases in which
modulation of the immune response is desired, e.g., for treating diseases related to
lymphocyte activation and migration, such as delayed type hypersensitivity reactions;
20   asthma; allergies; autoimmune diseases such as scleroderma, pernicious anemia, multiple
sclerosis, myasthenia gravis, IDDM, rheumatoid arthritis, systemic lupus erythematosus, and
Crohn's disease; and conditions related to organ and tissue transplant, such as graft vs. host
disease; and acute and chronic inflammation; as well as in diseases in which activation of the
immune response and stimulation of lymphocyte migration is desired, e.g., in
25   immunocompromised subjects, e.g., due to HIV infection or cancer; and in infectious disease
caused by viral, fungal, protozoal, and bacterial infections.

Definitions

By "disorder associated with lymphocyte activation or migration" or "disease
30   associated with lymphocyte activation or migration" herein is meant a disease state which is
marked by either an excess or a deficit of B or T cell activation or migration. For example,
lymphocyte activation disorders associated with increased activation or migration include,
but are not limited to, acute and chronic inflammation, asthma, allergies , autoimmune
disease and transplant rejection. Pathological states for which it may be desirable to increase

lymphocyte activation or migration include HIV infection that results in immunocompromise, cancer, and infectious disease such as viral, fungal, protozoal, and bacterial infections. Different compounds may be used to modulate lymphocyte activation and migration, or the same compound may be used to modulate lymphocyte activation and migration.

5   "Lymphocyte migration" refers to migration of B and T lymphocytes to and from primary and secondary lymphoid organs (e.g., bone marrow, thymus, lymph nodes, spleen, Peyer's patch, and tonsils), the periphery, and non-lymphoid tissues via the blood stream, lymphatic vessels, and by penetration of capillary walls (see, e.g., Paul, Immunology (3rd ed., 1993) (Chapters 4 and 6)).

10   "Lymphocyte activation" refers to the process of stimulating quiescent (G$_0$ phase of cell cycle), mature B and T cells by encounter with antigen, either directly or indirectly (e.g., via a helper cell and antigen presenting cells as well as via direct antigen contact with a cell surface molecule of the lymphocyte). Characteristics of activation can include, e.g., increase in cell surface markers such as CD69, entry into the G$_1$ phase of the

15   cell cycle, cytokine production, and proliferation (see, e.g., Paul, Immunology (3rd ed., 1993) (Chapters 13 and 14)).

          The terms "A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1,

20   TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH" protein or fragment thereof, or a nucleic acid encoding "A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B,

25   STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH" or a fragment thereof refer to nucleic acids and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%,

30   94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acids, to an amino acid sequence listed in Figure 7 and the sequence listing provided herein, or to an amino acid sequence encoded by a nucleic acid listed in Figure 7 and the sequence listing provided herein; (2) specifically bind to antibodies, e.g., polyclonal antibodies, raised against an

9

immunogen comprising an amino acid sequence of a protein listed in Figure 7 or the sequence listing, immunogenic fragments thereof, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to an anti-sense strand corresponding to a nucleic acid sequence encodingA-raf-1, Lck, Zap70, Syk, PLCγ1,

5   PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH, and conservatively modified variants thereof; (4) have a nucleic acid sequence that

10   has greater than about 60% sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%, or higher nucleotide sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, toA-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b,

15   6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH.

An A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1,

20   CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or any mammal. The nucleic acids

25   and proteins of the invention include both naturally occurring or recombinant molecules.

The Unigene number for EDG-1 is Hs.154210, and GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_001400, XM_001499, NP_001391, XP_00149, AAC51905, AAF43420, and AAA52336. The chromosomal location is Chr 1p21. The OMIM reference number for EDG-1 is 601974. EDG-1 is

30   expressed in, e.g., in endothelial cells, vascular smooth muscle cells, fibroblasts, melanocytes and cells of epithelioid origin (see, e.g., Hla & Maciag, J. Biol. Chem. 265:9308-9313 (1990); Hobson et al., Science 291:1800-1803 (2001); and Lee et al., Science 279:1552-1555 (1998)).

Exemplary wild type nucleic acid and protein sequences for additional members of the EDG family are provided by the following OMIM reference numbers (see also Figure 2 for exemplary amino acid sequences of EDG family members):

For EDG-2, OMIM reference number 602282. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_001401, XM_005557, XM_036690, XM_036691, NP_001392, XP)--5557, XP_036690, XP_036691, AAC00530, AAC51139, CAA70686, and CAA70687 (see, e.g., An et al., Molec. Pharm. 54:881-888 (1998); An et al., Biochem. Biophys. Res. Commun. 231:619-622 (1997); Contos et al., Genomics 51:364-378 (1998); Hecht et al., J. Cell. Biol. 135:1071-1083 (1996); and Moolenaar et al., Curr. Opin. Cell Biol. 9:168-173 (1997)).

For EDG-3, OMIM reference number 601965. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_005226, NP_005217, CAA58744 and AAC51906 (see, e.g., An et al., FEBS Lett. 417:279-282 (1997); and Yamaguchi et al., Biochem. Biophys. Res. Commun. 227:608-614 (1996)).

For EDG-4, OMIM reference number 605110. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_004720, XM_012893, XM_048494, XM_048495, NP_004711, XP_012893, XP_048494, XP_048495, AAB61528, AAC27728 and AAF43409 (see, e.g., An et al., J. Biol. Chem. 273:7906-7910 (1998); An et al., Molec. Pharm. 54:881-888 (1998); Contos et al., Genomics 64:155-169 (2000); and Goetzl et al., J. Immunol. 164:4996-4999 (2000)).

For EDG-5, OMIM reference number 605111. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_004230, XM_008898, NP_004221, XP_008898, and AAC98919 (see, e.g., An et al., J. Biol. Chem. 275:288-296 (2000); Kupperman et al., Nature 406:192-195 (2000); and MacLennan et al., Molec. Cell. Neurosci. 5:201-209 (1994)).

For EDG-6, OMIM reference number 603751. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_003775, XM_009219, NP_003766, XP_009219, and CAA04118 (see, e.g., Graler et al., Genomics 53:164-169 (1998); and Jedlicka et al., Cytogenet. Cell. Genet. 65:140 (1994)).

For EDG-7, OMIM reference number 605106. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_012152, XM_002057, XM_035234, NP_036284, XP_002057, XP_035234, AAD56311, AAF00530, and AAF91291 (see, e.g., Bandoh et al., J. Biol. Chem. 274:27776-27785 (1999)).

For EDG-8, OMIM reference number 605146. The GenBank accession numbers for exemplary nucleotide and amino acids sequences are NM_030760, XM_049584, NP_110387, XP_049584, and AAG3813 (*see, e.g.*, Im *et al., J. Biol. Chem.* 275:14281-14286 (2000)).

5   As described above, EDG proteins have "G-protein coupled receptor activity," e.g., they bind to G-proteins in response to extracellular stimuli, such as ligand binding, and promote production of second messengers such as IP3, cAMP, and Ca^{2+} via stimulation of enzymes such as phospholipase C and adenylate cyclase. Such activity can be measured in a heterologous cell, by coupling a GPCR (or a chimeric GPCR) to a G-protein, e.g., a

10   promiscuous G-protein such as $G\alpha15$, and an enzyme such as PLC, and measuring increases in intracellular calcium using (Offermans & Simon, *J. Biol. Chem.* 270:15175-15180 (1995)). Receptor activity can be effectively measured, e.g., by recording ligand-induced changes in $[Ca^{2+}]_i$ and calcium influx using fluorescent Ca^{2+}-indicator dyes and fluorometric imaging.

        G protein coupled receptors are glycoproteins that share certain structural

15   similarities (*see, e.g.*, Gilman, *Ann. Rev. Biochem.* 56:615-649 (1987), Strader *et al., The FASEB J.* 3:1825-1832 (1989), Kobilka *et al., Nature* 329:75-79 (1985), and Young *et al., Cell* 45:711-719 (1986)). For example, G protein coupled receptors have an extracellular domain, seven hydrophobic stretches of about 20-25 amino acids in length interspersed with eight hydrophilic regions (collectively known as the transmembrane domain), and a

20   cytoplasmic tail. Each of the seven hydrophobic regions forms a transmembrane alpha helix, with the intervening hydrophilic regions forming alternatively intracellular and extracellular loops. The third cytosolic loop between transmembrane domains five and six is involved in G-protein interaction. These transmembrane hydrophobic domains, hydrophilic loop domains, extracellular domains, and cytoplasmic tail domains can be structurally identified

25   using methods known to those of skill in the art, such as sequence analysis programs that identify hydrophobic and hydrophilic domains (*see, e.g.*, Kyte & Doolittle, *J. Mol. Biol.* 157:105-132 (1982)). Such domains are useful for making chimeric proteins and for *in vitro* assays of the invention (*see, e.g.*, WO 94/05695 and US Patent 5,508,384). Such domains are also considered "fragments" of EDG proteins, and as such are useful in the assays of the

30   invention, e.g., for ligand binding studies, or for signal transduction studies using chimeric proteins.

        The phrase "functional effects" in the context of assays for testing compounds that modulate activity of a protein listed in Figure 7 includes the determination of a parameter

that is indirectly or directly under the influence of a protein or nucleic acid listed in Figure 7, e.g., an indirect, chemical or phenotypic effect such as inhibition of lymphocyte activation or migration represented by a change in expression of a cell surface marker or cytokine production upon TCR stimulation, or changes in cellular proliferation or apoptosis, or signal

5   transduction leading to increases in intracellular calcium; or, e.g., a direct, physical effect such as ligand binding or inhibition of ligand binding or movement from one chamber to another in response to ligand. A functional effect therefore includes ligand binding activity, the ability of cells to proliferate, the ability of cells to migrate, apoptosis, gene expression in cells undergoing activation, expression of cell surface molecules such as CD69, signal

10   transduction, production of cytokines, calcium influx, and other characteristics of activated and/or migrating lymphocytes. "Functional effects" include *in vitro*, *in vivo*, and *ex vivo* activities.

By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a

15   protein listed in Figure 7, e.g., measuring physical and chemical or phenotypic effects. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein; measuring inducible markers or transcriptional activation of the protein; measuring binding activity or binding

20   assays, e.g. binding to antibodies; measuring changes in ligand binding affinity, e.g., SPP or LPA or analogs thereof or sphingolipid-like compounds, either naturally occurring or synthetic; measuring cellular proliferation; measuring cellular movement towards a ligand; measuring apoptosis; measuring cell surface marker expression, e.g., CD69; measuring cytokine, e.g., IL-2, production; measurement of calcium influx; measurement of changes in

25   protein levels for associated sequences; measurement of RNA stability; G-protein binding; GPCR phosphorylation or dephosphorylation; signal transduction, e.g., receptor-ligand interactions, second messenger concentrations (e.g., cAMP, IP3, or intracellular Ca^{2+}); identification of downstream or reporter gene expression (CAT, luciferase, β-gal, GFP and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding,

30   inducible markers, and ligand binding assays.

"Inhibitors", "activators", and "modulators" of polynucleotide and polypeptide sequences listed in Figure 7 are used to refer to activating, inhibitory, or modulating molecules identified using *in vitro* and *in vivo* assays of A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2,

EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH polynucleotide and polypeptide sequences. Inhibitors are

5   compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of these proteins, e.g., antagonists. "Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate protein activity. Inhibitors, activators, or modulators also include genetically modified versions of the proteins of Figure

10   7, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, peptides, cyclic peptides, nucleic acids, antibodies, antisense molecules, ribozymes, small organic molecules and the like. Such assays for inhibitors and activators include, e.g., expressing A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a,

15   PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH protein *in vitro*, in cells, cell extracts, or cell membranes, applying putative modulator compounds, and then determining the functional effects on activity, as described above.

20   Samples or assays comprising the proteins of Figure 7 that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of activation or migration modulation. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of a protein is achieved when the activity value relative to the control is

25   about 80%, preferably 50%, more preferably 25-0%. Activation of a protein is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

              The term "test compound" or "drug candidate" or "modulator" or grammatical

30   equivalents as used herein describes any molecule, either naturally occurring or synthetic, e.g., protein, oligopeptide (e.g., from about 5 to about 25 amino acids in length, preferably from about 10 to 20 or 12 to 18 amino acids in length, preferably 12, 15, or 18 amino acids in length), small organic molecule, polysaccharide, lipid, fatty acid, polynucleotide, oligonucleotide, etc., to be tested for the capacity to directly or indirectly modulation

14

lymphocyte activation. The test compound can be in the form of a library of test compounds, such as a combinatorial or randomized library that provides a sufficient range of diversity. Test compounds are optionally linked to a fusion partner, e.g., targeting compounds, rescue compounds, dimerization compounds, stabilizing compounds, addressable compounds, and

5   other functional moieties. Conventionally, new chemical entities with useful properties are generated by identifying a test compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

10   A "small organic molecule" refers to an organic molecule, either naturally occurring or synthetic, that has a molecular weight of more than about 50 daltons and less than about 2500 daltons, preferably less than about 2000 daltons, preferably between about 100 to about 1000 daltons, more preferably between about 200 to about 500 daltons.

                "Biological sample" include sections of tissues such as biopsy and autopsy
15   samples, and frozen sections taken for histologic purposes. Such samples include blood, sputum, tissue, cultured cells, e.g., primary cultures, explants, and transformed cells, stool, urine, etc. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

20   The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared

25   and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the compliment of a test sequence. The

30   definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

                                                              15

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default

5   program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20

10   to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl.*

15   *Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and

20   visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

A preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al., Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al., J.*

25   *Mol. Biol.* 215:403-410 (1990), respectively. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the

30   query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al., supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment

score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when:

5   the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation

10   (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

15   "Nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, and complements thereof. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are

20   metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

             Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions)

25   and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). The

30   term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

             A particular nucleic acid sequence also implicitly encompasses "splice variants." Similarly, a particular protein encoded by a nucleic acid implicitly encompasses any protein encoded by a splice variant of that nucleic acid. "Splice variants," as the name

suggests, are products of alternative splicing of a gene. After transcription, an initial nucleic acid transcript may be spliced such that different (alternate) nucleic acid splice products encode different polypeptides. Mechanisms for the production of splice variants vary, but include alternate splicing of exons. Alternate polypeptides derived from the same nucleic

5   acid by read-through transcription are also encompassed by this definition. Any products of a splicing reaction, including recombinant forms of the splice products, are included in this definition. An example of potassium channel splice variants is discussed in Leicher, *et al., J. Biol. Chem.* 273(52):35095-35101 (1998).

The terms "polypeptide," "peptide" and "protein" are used interchangeably

10   herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids,

15   as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is

20   bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an

25   amino acid, but that functions in a manner similar to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

30   "Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large

number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid

5   variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally

10   identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual probe sequences.

        As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein

15   sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude

20   polymorphic variants, interspecies homologs, and alleles of the invention.

        The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan

25   (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (see, e.g., Creighton, *Proteins* (1984)).

        Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, see, e.g., Alberts et al., *Molecular Biology of the Cell* (3rd ed., 1994) and Cantor & Schimmel,

30   *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains, e.g., transmembrane domains, pore domains, and cytoplasmic tail domains. Domains are portions of a polypeptide that form a compact

unit of the polypeptide and are typically 15 to 350 amino acids long. Exemplary domains include extracellular domains, transmembrane domains, and cytoplasmic domains. Typical domains are made up of sections of lesser organization such as stretches of β-sheet and α-helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ^{32}P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in

20

Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is

5   the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least

10   two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

Nucleic acids that do not hybridize to each other under stringent conditions are

15   still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl,

20   1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and *Current Protocols in Molecular Biology*, ed. Ausubel, *et al.*

25   For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for

30   both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification

reactions are provided, e.g., in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen.

5   The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody

10   will be most critical in specificity and affinity of binding.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily

15   responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'$_2$,

20   a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The F(ab)'$_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the F(ab)'$_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the digestion of an intact

25   antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (*see, e.g.*, McCafferty *et al.*, *Nature*

30   348:552-554 (1990))

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (*see, e.g.*, Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan,

Current Protocols in Immunology (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). The genes encoding the heavy and light chains of an antibody of interest can be cloned from a cell, e.g., the genes encoding a monoclonal antibody can be cloned from a hybridoma and

5 used to produce a recombinant monoclonal antibody. Gene libraries encoding heavy and light chains of monoclonal antibodies can also be made from hybridoma or plasma cells. Random combinations of the heavy and light chain gene products generate a large pool of antibodies with different antigenic specificity (*see, e.g.,* Kuby, *Immunology* (3rd ed. 1997)). Techniques for the production of single chain antibodies or recombinant antibodies (U.S.

10 Patent 4,946,778, U.S. Patent No. 4,816,567) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized or human antibodies (*see, e.g.,* U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, Marks *et al.,* *Bio/Technology* 10:779-783 (1992); Lonberg *et al., Nature* 368:856-859 (1994); Morrison,

15 *Nature* 368:812-13 (1994); Fishwild *et al., Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); and Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (*see, e.g.,* McCafferty *et al., Nature* 348:552-554 (1990); Marks *et al., Biotechnology* 10:779-783

20 (1992)). Antibodies can also be made bispecific, i.e., able to recognize two different antigens (*see, e.g.,* WO 93/08829, Traunecker *et al., EMBO J.* 10:3655-3659 (1991); and Suresh *et al., Methods in Enzymology* 121:210 (1986)). Antibodies can also be heteroconjugates, e.g., two covalently joined antibodies, or immunotoxins (*see, e.g.,* U.S. Patent No. 4,676,980 , WO 91/00360; WO 92/200373; and EP 03089).

25   Methods for humanizing or primatizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers

30 (*see, e.g.,* Jones *et al., Nature* 321:522-525 (1986); Riechmann *et al., Nature* 332:323-327 (1988); Verhoeyen *et al., Science* 239:1534-1536 (1988) and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human

variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

5   A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable

10   region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

In one embodiment, the antibody is conjugated to an "effector" moiety. The effector moiety can be any number of molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the

15   antibody modulates the activity of the protein.

The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, often in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions,

20   the specified antibodies bind to a particular protein at least two times the background and more typically more than 10 to 100 times background. Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a protein of Figure 7, polymorphic variants, alleles, orthologs, and conservatively modified variants, or splice variants, or portions

25   thereof, can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with a Figure 7 protein and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to

30   select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual* (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the

treatment, and will be ascertainable by one skilled in the art using known techniques (*see*, *e.g.*, Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); and Pickar, *Dosage Calculations* (1999)).

Assays for proteins that modulation lymphocyte activation

High throughput functional genomics assays can be used to identify modulators of lymphocyte activation. Such assays can monitor changes in cell surface marker expression, cytokine production, antibody production, proliferation and differentiation, and apoptosis, using either cell lines or primary cells. Typically, the lymphocytes are contacted with a cDNA or a random peptide library (encoded by nucleic acids). The cDNA library can comprise sense, antisense, full length, and truncated cDNAs. The peptide library is encoded by nucleic acids. The lymphocytes are then activated, e.g., by activating either the T cell receptor (TCR, also known as CD3) or the B cell receptor (BCR, also known as surface or mIg), as appropriate, e.g., using antibodies to the receptor. The effect of the cDNA or peptide library on the phenotype of lymphocyte activation is then monitored, using an assay as described above. The effect of the cDNA or peptide can be validated and distinguished from somatic mutations, using, e.g., regulatable expression of the nucleic acid such as expression from a tetracycline promoter. cDNAs and nucleic acids encoding peptides can be rescued using techniques known to those of skill in the art, e.g., using a sequence tag.

Proteins interacting with the peptide or with the protein encoded by the cDNA can be isolated using a yeast two-hybrid system, mammalian two hybrid system, or phage display screen, etc. Targets so identified can be further used as bait in these assays to identify additional members of the lymphocyte activation pathway, which members are also targets for drug development (*see, e.g.*, Fields *et al.*, *Nature* 340:245 (1989); Vasavada *et al.*, *Proc. Nat'l Acad. Sci. USA* 88:10686 (1991); Fearon *et al.*, *Proc. Nat'l Acad. Sci. USA* 89:7958 (1992); Dang *et al.*, *Mol. Cell. Biol.* 11:954 (1991); Chien *et al.*, *Proc. Nat'l Acad. Sci. USA* 9578 (1991); and U.S. Patent Nos. 5,283,173, 5,667,973, 5,468,614, 5,525,490, and 5,637,463).

Suitable B cell lines include surface Ig(+) lines such as CL-01, LA350, and CA46, as well as other mature and immature B cell lines and primary B cells known to those of skill in the art. Suitable T cell lines include Jurkat, HPB-ALL, HSB-2, and PEER, as well as other mature and immature T cell lines and primary T cells known to those of skill in the

25

art. Suitable B cell surface markers, for assaying B cell activation, include MHC class I, MHC class II, CD23, CD40, CD58, CD69, CD72, CD80, CD86, LFA-1, LFA-3, and ICAM-1, as well as other cell surface markers known to those of skill in the art. Suitable T cell surface markers include MHC class II, CD2, CD3, CD4, CD5, CD8, CD25, CD28, CD69,

5   CD40L, LFA-1, and ICAM-1 as well as other cell surface markers known to those of skill in the art (see, e.g., Yablonski et al., Science 281:413-416 (1998)). Suitable cytokines, for measuring either production or response, include IL-2, IL-4, IL-5, IL-6, IL-10, INF-γ, and TGF-β, as well as their corresponding receptors.

        Cell surface markers can be assayed using fluorescently labeled antibodies and

10   FACS. Cell proliferation can be measured using ^{3}H-thymidine or dye inclusion. Apoptosis can be measured using dye inclusion, or by assaying for DNA laddering or increases in intracellular calcium. Cytokine production can be measured using an immunoassay such as ELISA.

        cDNA libraries are made from any suitable source, preferably from primary

15   human lymphoid organs such as thymus, spleen, lymph node, and bone marrow. Libraries encoding random peptides are made according to techniques well known to those of skill in the art (see, e.g., U.S. Patent No. 6,153,380, 6,114,111, and 6,180,343). Any suitable vector can be used for the cDNA and peptide libraries, including, e.g., retroviral vectors.

        In a preferred embodiment, target proteins that modulate lymphocyte

20   activation, preferably T cell activation, are identified using a high throughput cell based assay (using a microtiter plate format) and FACS screening for CD69 cell surface expression (see Figures 1-4 and Example I). cDNA libraries are made from primary lymphocyte organs. These cDNA libraries include, e.g., sense, antisense, full length, and truncated cDNAs. The cDNAs are cloned into a retroviral vector with a tet-regulatable promoter. Jurkat cells are

25   infected with the library, the cells are stimulated with anti-TCR antibodies, and then the cells are sorted using fluorescent antibodies and FACS for CD69 low/CD3+ cells. Cells with the desired phenotype are recovered, expanded, and cloned. A Tet-regulatable phenotype is established to distinguish somatic mutations. The cDNA is rescued. Optionally, the phenotype is validated by assaying for IL-2 production using primary lymphocytes.

30   Optionally, a marker such as GFP can be used to select for retrovirally infected cells. Using this system, cDNAs encoding the proteins of Figure 1 were identified as inhibitors of T cell activation.

Isolation of nucleic acids

This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989); Kriegler, *Gene*

5   *Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994)).

Nucleic acids, polymorphic variants, orthologs, and alleles that are substantially identical to an amino acid sequence encoded by a sequence in Figure 7 or the sequence listing, can be isolated using nucleic acid probes and oligonucleotides under

10   stringent hybridization conditions, by screening libraries. Alternatively, expression libraries can be used to clone a protein, polymorphic variants, orthologs, and alleles by detecting expressed homologs immunologically with antisera or purified antibodies made against a human protein or portions thereof.

To make a cDNA library, one should choose a source that is rich in the

15   selected RNA. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known (*see, e.g.*, Gubler & Hoffman, *Gene* 25:263-269 (1983); Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*).

For a genomic library, the DNA is extracted from the tissue and either

20   mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, *Science* 196:180-182 (1977). Colony hybridization is carried out as generally described in

25   Grunstein *et al.*, *Proc. Natl. Acad. Sci. USA.*, 72:3961-3965 (1975).

An alternative method of isolating nucleic acids and orthologs, alleles, mutants, polymorphic variants, and conservatively modified variants combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (*see* U.S. Patents 4,683,195 and 4,683,202; *PCR Protocols: A Guide to Methods and Applications*

30   (Innis *et al.*, eds, 1990)). Methods such as polymerase chain reaction (PCR) and ligase chain reaction (LCR) can be used to amplify nucleic acid sequences directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify protein homologs using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other

in vitro amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of a selected mRNA in physiological samples, for nucleic acid sequencing, or for other purposes. Genes amplified by the PCR reaction can be purified from

5   agarose gels and cloned into an appropriate vector.

Gene expression can also be analyzed by techniques known in the art, e.g., reverse transcription and amplification of mRNA, isolation of total RNA or poly A^+ RNA, northern blotting, dot blotting, *in situ* hybridization, RNase protection, high density polynucleotide array technology, e.g., and the like.

10   Nucleic acids can be used with high density oligonucleotide array technology (e.g., GeneChip™) to identify proteins, orthologs, alleles, conservatively modified variants, and polymorphic variants in this invention. In the case where the homologs being identified are linked to modulation of T cell activation and migration, they can be used with GeneChip™ as a diagnostic tool in detecting the disease in a biological sample, *see, e.g.,*

15   Gunthand *et al., AIDS Res. Hum. Retroviruses* 14: 869-876 (1998); Kozal *et al., Nat. Med.* 2:753-759 (1996); Matson *et al., Anal. Biochem.* 224:110-106 (1995); Lockhart *et al., Nat. Biotechnol.* 14:1675-1680 (1996); Gingeras *et al., Genome Res.* 8:435-448 (1998); Hacia *et al., Nucleic Acids Res.* 26:3865-3866 (1998).

The selected gene is typically cloned into intermediate vectors before

20   transformation into prokaryotic or eukaryotic cells for replication and/or expression. These intermediate vectors are typically prokaryote vectors, e.g., plasmids, or shuttle vectors.

Expression in prokaryotes and eukaryotes

To obtain high level expression of a cloned gene, one typically subclones the

25   nucleic acid into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, e.g., in Sambrook *et al.,* and Ausubel *et al, supra.* Bacterial expression systems for expressing the protein are available in, e.g., *E. coli, Bacillus sp.,* and *Salmonella*

30   (Palva *et al., Gene* 22:229-235 (1983); Mosbach *et al., Nature* 302:543-545 (1983). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. In one preferred embodiment, retroviral expression systems are used in the present invention.

Selection of the promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can

5   be accommodated without loss of promoter function.

In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the nucleic acid in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence and signals required for efficient

10   polyadenylation of the transcript, ribosome binding sites, and translation termination. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient

15   termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include

20   plasmids such as pBR322 based plasmids, pSKF, pET23D, and fusion expression systems such as MBP, GST, and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, e.g., c-myc. Sequence tags may be included in an expression cassette for nucleic acid rescue. Markers such as fluorescent proteins, green or red fluorescent protein, β-gal, CAT, and the like can be included in the vectors as markers for

25   vector transduction.

Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, e.g., SV40 vectors, papilloma virus vectors, retroviral vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A$^+$, pMTO10/A$^+$, pMAMneo-5, baculovirus pDSVE, and

30   any other vector allowing expression of proteins under the direction of the CMV promoter, SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

29

Expression of proteins from eukaryotic vectors can be also be regulated using inducible promoters. With inducible promoters, expression levels are tied to the concentration of inducing agents, such as tetracycline or ecdysone, by the incorporation of response elements for these agents into the promoter. Generally, high level expression is

5   obtained from inducible promoters only in the presence of the inducing agent; basal expression levels are minimal.

In one embodiment, the vectors of the invention have a regulatable promoter, e.g., tet-regulated systems and the RU-486 system (see, e.g., Gossen & Bujard, PNAS 89:5547 (1992); Oligino et al., Gene Ther. 5:491-496 (1998); Wang et al., Gene Ther. 4:432-

10   441 (1997); Neering et al., Blood 88:1147-1155 (1996); and Rendahl et al., Nat. Biotechnol. 16:757-761 (1998)). These impart small molecule control on the expression of the candidate target nucleic acids. This beneficial feature can be used to determine that a desired phenotype is caused by a transfected cDNA rather than a somatic mutation.

Some expression systems have markers that provide gene amplification such

15   as thymidine kinase and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a protein encoding sequence under the direction of the polyhedrin promoter or other strong baculovirus promoters.

The elements that are typically included in expression vectors also include a

20   replicon that functions in E. coli, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are preferably chosen such that they do not interfere with the

25   replication of the DNA in eukaryotic cells, if necessary.

Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of the protein, which are then purified using standard techniques (see, e.g., Colley et al., J. Biol. Chem. 264:17619-17622 (1989); Guide to Protein Purification, in Methods in Enzymology, vol. 182 (Deutscher, ed., 1990)).

30   Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (see, e.g., Morrison, J. Bact. 132:349-351 (1977); Clark-Curtiss & Curtiss, Methods in Enzymology 101:347-362 (Wu et al., eds, 1983).

Any of the well-known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate

transfection, polybrene, protoplast fusion, electroporation, biolistics, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Sambrook *et al., supra*). It is only necessary that the particular

5   genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing the protein.

After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of the protein, which is recovered from the culture using standard techniques identified below.

10

PURIFICATION OF POLYPEPTIDES

Either naturally occurring or recombinant protein can be purified for use in functional assays. Naturally occurring protein can be purified, e.g., from human tissue. Recombinant protein can be purified from any suitable expression system.

15   The protein may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (*see, e.g.*, Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al., supra;* and Sambrook *et al., supra*).

20   A number of procedures can be employed when recombinant protein is being purified. For example, proteins having established molecular adhesion properties can be reversible fused to the protein. With the appropriate ligand, the protein can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally, EDG protein could be

25   purified using immunoaffinity columns.

A.   Purification of protein from recombinant bacteria

Recombinant proteins are expressed by transformed bacteria in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction

30   with IPTG is one example of an inducible promoter system. Bacteria are grown according to standard procedures in the art. Fresh or frozen bacteria cells are used for isolation of protein.

Proteins expressed in bacteria may form insoluble aggregates ("inclusion bodies"). Several protocols are suitable for purification of protein inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or

31

purification of inclusion bodies by disruption of bacterial cells, e.g., by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM MgCl$_2$, 1 mM DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension can be lysed using 2-3 passages through a French Press, homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate

5   methods of lysing bacteria are apparent to those of skill in the art (*see, e.g.*, Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*).

If necessary, the inclusion bodies are solubilized, and the lysed cell suspension is typically centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable

10   solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, for example SDS (sodium dodecyl sulfate), 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a

15   lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art. Human proteins are separated from other bacterial proteins by

20   standard separation techniques, e.g., with Ni-NTA agarose resin.

Alternatively, it is possible to purify protein from bacteria periplasm. After lysis of the bacteria, when the protein is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to skill in the art. To isolate recombinant proteins from the periplasm, the

25   bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM MgSO$_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard

30   separation techniques well known to those of skill in the art.

B.   Standard protein separation techniques for purifying proteins

<u>Solubility fractionation</u>

Often as an initial step, particularly if the protein mixture is complex, an initial

salt fractionation can separate many of the unwanted host cell proteins (or proteins derived

5   from the cell culture media) from the recombinant protein of interest. The preferred salt is

ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the

amount of water in the protein mixture. Proteins then precipitate on the basis of their

solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower

ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium

10   sulfate to a protein solution so that the resultant ammonium sulfate concentration is between

20-30%. This concentration will precipitate the most hydrophobic of proteins. The

precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium

sulfate is added to the supernatant to a concentration known to precipitate the protein of

interest. The precipitate is then solubilized in buffer and the excess salt removed if

15   necessary, either through dialysis or diafiltration. Other methods that rely on solubility of

proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can

be used to fractionate complex protein mixtures.

<u>Size differential filtration</u>

20   The molecular weight of the protein can be used to isolate it from proteins of

greater and lesser size using ultrafiltration through membranes of different pore size (for

example, Amicon or Millipore membranes). As a first step, the protein mixture is

ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off

than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then

25   ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of

the protein of interest. The recombinant protein will pass through the membrane into the

filtrate. The filtrate can then be chromatographed as described below.

<u>Column chromatography</u>

30   The protein can also be separated from other proteins on the basis of its size,

net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised

against proteins can be conjugated to column matrices and the proteins immunopurified. All

of these methods are well known in the art. It will be apparent to one of skill that

33

chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

ASSAYS FOR MODULATORS OF PROTEINS INVOLVED IN LYMPHOCYTE

5   ## ACTIVATION

   A.   Assays

   Modulation of a protein as listed in Figure 7, and corresponding modulation of lymphocyte activation and/or migration, can be assessed using a variety of *in vitro* and *in vivo* assays, including cell-based models as described above. Such assays can be used to test for

10   inhibitors and activators of A-raf-1, Lck, Zap70, Syk, PLCγ1, PAG, SHP/PTP1C, CSK, nucleolin, SLAP, PAK2, TRAC1, TCPTP/PTPN2, EDG1, IL10-Rα, integrinα2, Enolase 1a, PRSM1, CLN2, P2X5b, 6PFKL, DUSP1, KIAA0251, GG2-1, GRB7, SH2-B, STAT1, TCF19, HFP101S, RERE, SudD, Ku70, SCAMP2, Fibulin-5, KIAA1228, Est from clone 2108068, vimentin, filamin A α, centractin α, moesin, TIMP3, and RNH protein or fragments

15   thereof, and, consequently, inhibitors and activators of lymphocyte activation and migration. Such modulators are useful for treating disorders related to T and B cell activation and migration. Modulators are tested using either recombinant or naturally occurring protein, preferably human protein.

   Preferably, the protein will have the sequence as listed in the sequence listing

20   provided herein, or in an application incorporated by reference, or an exemplary Genbank Accession number as provided herein (*see, e.g.*, Figure 7), or a conservatively modified variant thereof. Alternatively, the protein of the assay will be derived from a eukaryote and include an amino acid subsequence having substantial amino acid sequence identity to a sequence listed herein. Generally, the amino acid sequence identity will be at least 60%,

25   preferably at least 65%, 70%, 75%, 80%, 85%, or 90%, most preferably at least 95%.

   Measurement of lymphocyte activation, migration, or loss-of-lymphocyte activation or migration phenotype of the protein or cell expressing the protein, either recombinant or naturally occurring, can be performed using a variety of assays, *in vitro, in vivo*, and *ex vivo*, as described herein. A suitable physical, chemical or phenotypic change

30   that affects activity or binding can be used to assess the influence of a test compound on the polypeptide of this invention. When the functional effects are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of signal transduction, e.g., ligand binding (SPP, LPA, GTP), hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), cellular movement

towards a ligand, movement of labeled cells, changes in cell metabolism such as pH changes, and changes in intracellular second messengers such as Ca^{2+}, IP3, cGMP, or cAMP; as well as changes related to lymphocyte activation and migration, e.g., cellular proliferation, cell surface marker expression, e.g., CD69, cytokine production, and apoptosis.

5   In one preferred embodiment, described herein in Example I, measurement of CD69 activation and FACS sorting is used to identify modulators of lymphocyte, e.g., T cell, activation. In another preferred embodiment, measurement of cellular migration toward a ligand is used to identify modulators of lymphocyte, e.g., T cell, migration.

10   *In vitro* assays

         Assays to identify compounds with lymphocyte activation-modulating activity can be performed *in vitro*. Such assays can used full length protein or a variant thereof, or a fragment of a protein, such as an extracellular domain or a cytoplasmic domain, optionally fused to a heterologous protein to form a chimera. In one embodiment, different domains can

15   be used to assay for activation and migration. In another embodiment, the same domain can be used to assay for activation and migration. Purified recombinant or naturally occurring protein can be used in the *in vitro* methods of the invention. In addition to purified protein or fragment thereof, the recombinant or naturally occurring protein can be part of a cellular lysate or a cell membrane. As described below, the binding assay can be either solid state or

20   soluble. Preferably, the protein, fragment thereof or membrane is bound to a solid support, either covalently or non-covalently. Often, the *in vitro* assays of the invention are ligand binding or ligand affinity assays, either non-competitive or competitive (with known extracellular ligands SPP or LPA, or with a known intracellular ligand GTP). Other *in vitro* assays include measuring changes in spectroscopic (e.g., fluorescence, absorbance, refractive

25   index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein.

         In one embodiment, a high throughput binding assay is performed in which the protein or fragment thereof is contacted with a potential modulator and incubated for a suitable amount of time. In one embodiment, the potential modulator is bound to a solid support, and the protein is added. In another embodiment, the protein is bound to a solid

30   support. A wide variety of modulators can be used, as described below, including small organic molecules, peptides, antibodies, and ligand analogs. A wide variety of assays can be used to identify modulator binding, including labeled protein-protein binding assays, electrophoretic mobility shifts, immunoassays, enzymatic assays such as phosphorylation assays, and the like. In some cases, the binding of the candidate modulator is determined

through the use of competitive binding assays, where interference with binding of a known ligand is measured in the presence of a potential modulator. Ligands for the EDG family are known (SPP, LPA, and GTP). Either the modulator or the known ligand is bound first, and then the competitor is added. After the protein is washed, interference with binding, either of

5   the potential modulator or of the known ligand, is determined. Often, either the potential modulator or the known ligand is labeled.

Cell-based *in vivo* assays

In another embodiment, the protein is expressed in a cell, and functional, e.g.,

10   physical and chemical or phenotypic, changes are assayed to identify the protein and lymphocyte activation and migration modulators. Cells expressing the proteins of the invention can also be used in binding assays. Any suitable functional effect can be measured, as described herein. For example, ligand binding, cell surface marker expression, cellular proliferation, apoptosis, cytokine production, and GPCR signal transduction, e.g., changes in

15   intracellular Ca^{2+} levels, are all suitable assays to identify potential modulators using a cell based system. Suitable cells for such cell based assays include both primary lymphocytes and cell lines, as described herein. The protein can be naturally occurring or recombinant. Also, as described above, fragments of proteins or chimeras can be used in cell based assays.

As described above, in one embodiment, lymphocyte activation is measured

20   by contacting T cells comprising a target protein with a potential modulator and activating the cells with an anti-TCR antibody. Modulation of T cell activation is identified by screening for cell surface marker expression, e.g., CD69 expression levels, using fluorescent antibodies and FACS sorting. In another embodiment, lymphocyte migration is measured by observing T cell migration from an upper to a lower chamber containing a ligand.

25   In another embodiment, cellular proliferation, migration, or apoptosis can be measured using ^{3}H-thymidine incorporation or dye inclusion. Cytokine production can be measured using an immunoassay such as an ELISA.

In another embodiment, cellular polypeptide levels are determined by measuring the level of protein or mRNA. The level of protein or proteins are measured using

30   immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNAse protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled

nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described
herein.

Alternatively, protein expression can be measured using a reporter gene
system. Such a system can be devised using a protein promoter operably linked to a reporter
5   gene such as chloramphenicol acetyltransferase, firefly luciferase, bacterial luciferase, β-
galactosidase and alkaline phosphatase. Furthermore, the protein of interest can be used as an
indirect reporter via attachment to a second reporter such as red or green fluorescent protein
(see, e.g., Mistili & Spector, Nature Biotechnology 15:961-964 (1997)). The reporter
construct is typically transfected into a cell. After treatment with a potential modulator, the
10   amount of reporter gene transcription, translation, or activity is measured according to
standard techniques known to those of skill in the art.

In another embodiment, a functional effect related to GPCR signal
transduction can be measured. An activated or inhibited G-coupled protein receptor will alter
the properties of target enzymes, second messengers, channels, and other effector proteins.
15   The examples include the activation of cGMP phosphodiesterase, adenylate cyclase,
phospholipase C, IP3, and modulation of diverse channels by G proteins. Downstream
consequences can also be examined such as generation of diacyl glycerol and IP3 by
phospholipase C, and in turn, for calcium mobilization by IP3. Activated GPCR receptors
become substrates for kinases that phosphorylate the C-terminal tail of the receptor (and
20   possibly other sites as well). Thus, activators will promote the transfer of ^{32}P from gamma-
labeled GTP to the receptor, which can be assayed with a scintillation counter. The
phosphorylation of the C-terminal tail will promote the binding of arrestin-like proteins and
will interfere with the binding of G-proteins. For a general review of GPCR signal
transduction and methods of assaying signal transduction, see, e.g., Methods in Enzymology,
25   vols. 237 and 238 (1994) and volume 96 (1983); Bourne et al., Nature 10:349:117-27 (1991);
Bourne et al., Nature 348:125-32 (1990); Pitcher et al., Annu. Rev. Biochem. 67:653-92
(1998).

As described above, activation of some G-protein coupled receptors stimulates
the formation of inositol triphosphate (IP3) through phospholipase C-mediated hydrolysis of
30   phosphatidylinositol (Berridge & Irvine, Nature 312:315-21 (1984)). IP3 in turn stimulates
the release of intracellular calcium ion stores. Thus, a change in cytoplasmic calcium ion
levels, or a change in second messenger levels such as IP3 can be used to assess G-protein
coupled receptor function. Cells expressing such G-protein coupled receptors may exhibit

increased cytoplasmic calcium levels as a result of contribution from both intracellular stores and via activation of ion channels, in which case it may be desirable although not necessary to conduct such assays in calcium-free buffer, optionally supplemented with a chelating agent such as EGTA, to distinguish fluorescence response resulting from calcium release from

5   internal stores.

In one example, GPCR activity is measured by expressing a GPCR in a heterologous cell with a promiscuous G-protein that links the receptor to a phospholipase C signal transduction pathway (*see* Offermanns & Simon, *J. Biol. Chem.* 270:15175-15180 (1995)). Modulation of signal transduction is assayed by measuring changes in intracellular

10 Ca^{2+} levels, which change in response to modulation of the GPCR signal transduction pathway via administration of a molecule that associates with a GPCR. Changes in Ca^{2+} levels are optionally measured using fluorescent Ca^{2+} indicator dyes and fluorometric imaging.

In another example, phosphatidyl inositol (PI) hydrolysis can be analyzed

15 according to U.S. Patent 5,436,128, herein incorporated by reference. Briefly, the assay involves labeling of cells with ^{3}H-myoinositol for 48 or more hrs. The labeled cells are treated with a test compound for one hour. The treated cells are lysed and extracted in chloroform-methanol-water after which the inositol phosphates were separated by ion exchange chromatography and quantified by scintillation counting. Fold stimulation is

20 determined by calculating the ratio of cpm in the presence of agonist to cpm in the presence of buffer control. Likewise, fold inhibition is determined by calculating the ratio of cpm in the presence of antagonist to cpm in the presence of buffer control (which may or may not contain an agonist).

Other assays can involve determining the activity of receptors which, when

25 activated, result in a change in the level of intracellular cyclic nucleotides, e.g., cAMP or cGMP, by activating or inhibiting enzymes such as adenylate cyclase. In cases where activation of the receptor results in a decrease in cyclic nucleotide levels, it may be preferable to expose the cells to agents that increase intracellular cyclic nucleotide levels, e.g., forskolin, prior to adding a receptor-activating compound to the cells in the assay.

30 In one example, the changes in intracellular cAMP or cGMP can be measured using immunoassays. The method described in Offermanns & Simon, *J. Biol. Chem.* 270:15175-15180 (1995) may be used to determine the level of cAMP. Also, the method described in Felley-Bosco *et al.*, *Am. J. Resp. Cell and Mol. Biol.* 11:159-164 (1994) may be

used to determine the level of cGMP. Further, an assay kit for measuring cAMP and/or cGMP is described in U.S. Patent 4,115,538, herein incorporated by reference.

In one example, assays for G-protein coupled receptor activity include cells that are loaded with ion or voltage sensitive dyes to report receptor activity, e.g., by

5   observing calcium influx or intracellular calcium release. Assays for determining activity of such receptors can also use known agonists and antagonists for other G-protein coupled receptors as negative or positive controls to assess activity of tested compounds. In assays for identifying modulatory compounds (e.g., agonists, antagonists), changes in the level of ions in the cytoplasm or membrane voltage will be monitored using an ion sensitive or

10 membrane voltage fluorescent indicator, respectively. Among the ion-sensitive indicators and voltage probes that may be employed are those disclosed in the Molecular Probes 1997 Catalog. For G-protein coupled receptors, promiscuous G-proteins such as Gα15 and Gα16 can be used in the assay of choice (Wilkie *et al., Proc. Nat'l Acad. Sci. USA* 88:10049-10053 (1991)). Such promiscuous G-proteins allow coupling of a wide range of receptors.

15

Animal models

Animal models of lymphocyte activation and migration also find use in screening for modulators of lymphocyte activation or migration. Similarly, transgenic animal technology including gene knockout technology, for example as a result of homologous

20 recombination with an appropriate gene targeting vector, or gene overexpression, will result in the absence or increased expression of the protein. The same technology can also be applied to make knock-out cells. When desired, tissue-specific expression or knockout of the protein may be necessary. Transgenic animals generated by such methods find use as animal models of lymphocyte activation and migration and are additionally useful in screening for

25 modulators of lymphocyte activation and migration.

Knock-out cells and transgenic mice can be made by insertion of a marker gene or other heterologous gene into an endogenous gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting an endogenous gene with a mutated version of the gene, or by mutating an endogenous gene, e.g., by exposure to

30 carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice

that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (see, e.g., Capecchi et al., Science 244:1288 (1989)). Chimeric targeted mice can be derived according to Hogan et al., Manipulating the Mouse Embryo: A Laboratory Manual,

5   Cold Spring Harbor Laboratory (1988) and Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, ed., IRL Press, Washington, D.C., (1987).

        B.   Modulators
                The compounds tested as modulators of a protein can be any small organic

10   molecule, or a biological entity, such as a protein, e.g., an antibody or peptide, a sugar, a nucleic acid, e.g., an antisense oligonucleotide or a ribozyme, or a lipid. Alternatively, modulators can be genetically altered versions of a protein. Typically, test compounds will be small organic molecules, peptides, lipids, and lipid analogs. In one embodiment, the compound is a sphingolipid analog, either naturally occurring or synthetic. In another

15   embodiment, the compound is 2-amino-2(2-[4-octylphenyl]ethyl)-1,3-propanediol hydrochloride (also known as FTY720) or an analog thereof.
                Essentially any chemical compound can be used as a potential modulator or ligand in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to

20   screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs Switzerland) and the

25   like.
                In one preferred embodiment, high throughput screening methods involve providing a combinatorial small organic molecule or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such "combinatorial chemical libraries" or "ligand libraries" are then screened in one or

30   more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical

5   building blocks (amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to,

10   peptide libraries (see, e.g., U.S. Patent 5,010,175, Furka, Int. J. Pept. Prot. Res. 37:487-493 (1991) and Houghton et al., Nature 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (e.g., PCT Publication No. WO 91/19735), encoded peptides (e.g., PCT Publication No. WO 93/20242), random bio-oligomers (e.g., PCT Publication No. WO 92/00091),

15   benzodiazepines (e.g., U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs et al., Proc. Nat. Acad. Sci. USA 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara et al., J. Amer. Chem. Soc. 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann et al., J. Amer. Chem. Soc. 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen

20   et al., J. Amer. Chem. Soc. 116:2661 (1994)), oligocarbamates (Cho et al., Science 261:1303 (1993)), and/or peptidyl phosphonates (Campbell et al., J. Org. Chem. 59:658 (1994)), nucleic acid libraries (see Ausubel, Berger and Sambrook, all supra), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), antibody libraries (see, e.g., Vaughn et al., Nature Biotechnology, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (see,

25   e.g., Liang et al., Science, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (see, e.g., benzodiazepines, Baum C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

30   Devices for the preparation of combinatorial libraries are commercially available (see, e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (see, e.g., ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis,

MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

C.   Solid state and soluble high throughput assays

5   In one embodiment the invention provides soluble assays using a protein, or a cell or tissue expressing a protein, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the protein or fragment thereof, such as the cytoplasmic domain, is attached to a solid phase substrate. Any one of the assays described herein can be adapted for high

10   throughput screening, e.g., ligand binding, cellular proliferation, cell surface marker flux, e.g., CD-69, screening, radiolabeled GTP binding, second messenger flux, e.g., Ca^{2+}, IP3, cGMP, or cAMP, cytokine production, etc. In one preferred embodiment, the cell-based system using CD-69 modulation and FACS assays is used in a high throughput format for   · identifying modulators of Figure 7 proteins, and therefore modulators of T cell activation.

15   In the high throughput assays of the invention, either soluble or solid state, it is possible to screen up to several thousand different modulators or ligands in a single day. This methodology can be used for proteins *in vitro*, or for cell-based or membrane-based assays comprising a protein of Figure 7. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation

20   time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay many plates per day; assay screens for up to about 6,000, 20,000, 50,000, or more than 100,000 different compounds are possible using the integrated systems of the

25   invention.

          For a solid state reaction, the protein of interest or a fragment thereof, e.g., an extracellular domain, or a cell or membrane comprising the protein of interest or a fragment thereof as part of a fusion protein can be bound to the solid state component, directly or indirectly, via covalent or non covalent linkage e.g., via a tag. The tag can be any of a variety

30   of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest is attached to the solid support by interaction of the tag and the tag binder.

          A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder,

for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag

binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, etc.)

Antibodies to molecules with natural binders such as biotin are also widely available and

appropriate tag binders; *see,* SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis

5   MO).

Similarly, any haptenic or antigenic compound can be used in combination

with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies

are commercially available and many additional antibodies are described in the literature.

For example, in one common configuration, the tag is a first antibody and the tag binder is a

10   second antibody which recognizes the first antibody. In addition to antibody-antigen

interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs. For

example, agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand

interactions such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine

receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherein

15   family, the integrin family, the selectin family, and the like; *see, e.g.,* Pigott & Power, *The*

Adhesion Molecule Facts Book I (1993). Similarly, toxins and venoms, viral epitopes,

hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g. which mediate the effects

of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D;

peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer

20   configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with

various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates,

polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides,

and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder

25   pairs are also useful in assay systems described herein, as would be apparent to one of skill

upon review of this disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as

tags, and include polypeptide sequences, such as poly gly sequences of between about 5 and

200 amino acids. Such flexible linkers are known to persons of skill in the art. For example,

30   poly(ethelyne glycol) linkers are available from Shearwater Polymers, Inc. Huntsville,

Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or

heterofunctional linkages.

Tag binders are fixed to solid substrates using any of a variety of methods

currently available. Solid substrates are commonly derivatized or functionalized by exposing

43

all or a portion of the substrate to a chemical reagent which fixes a chemical group to the

surface which is reactive with a portion of the tag binder. For example, groups which are

suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and

carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a

5   variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer

arrays is well described in the literature. *See, e.g.,* Merrifield, *J. Am. Chem. Soc.* 85:2149-

2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen *et al., J. Immun.

Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank &

Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences

10   on cellulose disks); Fodor *et al., Science,* 251:767-777 (1991); Sheldon *et al., Clinical

Chemistry* 39(4):718-719 (1993); and Kozal *et al., Nature Medicine* 2(7):753759 (1996) (all

describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for

fixing tag binders to substrates include other common methods, such as heat, cross-linking by

UV radiation, and the like.

15

IMMUNOLOGICAL DETECTION OF POLYPEPTIDES

            In addition to the detection of gene and gene expression using nucleic acid

hybridization technology, one can also use immunoassays to detect proteins of the invention.

Such assays are useful for screening for modulators of lymphocyte activation and migration,

20   as well as for therapeutic and diagnostic applications. Immunoassays can be used to

qualitatively or quantitatively analyze a selected protein. A general overview of the

applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual*

(1988).

25   *A.   Production of antibodies*

            Methods of producing polyclonal and monoclonal antibodies that react

specifically with the selected proteins are known to those of skill in the art (*see, e.g.,* Coligan,

Current Protocols in Immunology (1991); Harlow & Lane, *supra*; Goding, *Monoclonal

Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-

30   497 (1975). Such techniques include antibody preparation by selection of antibodies from

libraries of recombinant antibodies in phage or similar vectors, as well as preparation of

polyclonal and monoclonal antibodies by immunizing rabbits or mice (*see, e.g.,* Huse *et al.,

Science* 246:1275-1281 (1989); Ward *et al., Nature* 341:544-546 (1989)).

A number of immunogens comprising portions of the selected protein may be used to produce antibodies specifically reactive with the protein. For example, recombinant protein or an antigenic fragment thereof, can be isolated as described herein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified

5 as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either

10 monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization

15 protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the beta subunits. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (see, Harlow & Lane, supra).

20 Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see, Kohler & Milstein, Eur. J. Immunol. 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well

25 known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof

30 by screening a DNA library from human B cells according to the general protocol outlined by Huse, et al., Science 246:1275-1281 (1989).

Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4

or greater are selected and tested for their cross reactivity against non-immunogen proteins, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better.

5   Antibodies specific only for a particular family member, or a particular ortholog, such as human protein, can also be made, by subtracting out other cross-reacting family members or orthologs from a species such as a non-human mammal. In this manner, antibodies that bind only to a particular protein or ortholog may be obtained.

Once the specific antibodies against a selected protein are available, the

10   protein can be detected by a variety of immunoassay methods. In addition, the antibody can be used therapeutically as a lymphocyte activation modulators. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in Enzyme

15   Immunoassay (Maggio, ed., 1980); and Harlow & Lane, *supra*.

B.   *Immunological binding assays*

Protein can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.,* U.S. Patents 4,366,241; 4,376,110;

20   4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991). Immunological binding assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice. The antibody may be produced by any of a number of means well known to those of

25   skill in the art and as described above.

Immunoassays also often use a labeling agent to specifically bind to and label the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled protein or a labeled antibody. Alternatively, the labeling agent may be a third moiety,

30   such a secondary antibody, that specifically binds to the antibody/protein complex (a secondary antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a

variety of species (*see, e.g.,* Kronval *et al., J. Immunol.* 111:1401-1406 (1973); Akerstrom *et al., J. Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

5   Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, optionally from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be

10   conducted over a range of temperatures, such as 10°C to 40°C.

Non-competitive assay formats

Immunoassays for detecting a selected protein in samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the

15   amount of antigen is directly measured. In one preferred "sandwich" assay, for example, the anti-immunogen antibodies can be bound directly to a solid substrate on which they are immobilized. These immobilized antibodies then capture the immunogen present in the test sample. Proteins thus immobilized are then bound by a labeling agent, such as a second anti-immunogen antibody bearing a label. Alternatively, the second antibody may lack a label,

20   but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibody is typically modified with a detectable moiety, such as biotin, to which another molecule specifically binds, e.g., streptavidin, to provide a detectable moiety.

25   ## Competitive assay formats

In competitive assays, the amount of a selected protein present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) protein displaced (competed away) from an anti-immunogen antibody by the unknown immunogen protein present in a sample. In one competitive assay, a known amount of immunogen

30   protein is added to a sample and the sample is then contacted with an antibody that specifically binds to a selected protein. The amount of exogenous protein bound to the antibody is inversely proportional to the concentration of protein present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. The

amount of immunogen protein bound to the antibody may be determined either by measuring the amount of immunogen present in immunogen protein/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of immunogen protein may be detected by providing a labeled immunogen molecule.

5   A hapten inhibition assay is another preferred competitive assay. In this assay the known immunogen protein is immobilized on a solid substrate. A known amount of anti-immunogen antibody is added to the sample, and the sample is then contacted with the immobilized immunogen. The amount of anti-immunogen antibody bound to the known immobilized immunogen is inversely proportional to the amount of immunogen protein

10   present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

15

Cross-reactivity determinations

          Immunoassays in the competitive binding format can also be used for crossreactivity determinations. For example, a selected immunogen protein can be immobilized to a solid support. Proteins are added to the assay that compete for binding of

20   the antisera to the immobilized antigen. The ability of the added proteins to compete for binding of the antisera to the immobilized protein is compared to the ability of the immunogen protein to compete with itself. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the added proteins listed above are selected and pooled. The cross-reacting

25   antibodies are optionally removed from the pooled antisera by immunoabsorption with the added considered proteins, e.g., distantly related homologs.

          The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps an allele or polymorphic variant of the selected protein, to the immunogen protein. In order

30   to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required to inhibit 50% of binding is less than 10 times the amount of the immunogen protein that is required to

inhibit 50% of binding, then the second protein is said to specifically bind to the polyclonal antibodies generated to the immunogen.

Other assay formats

5   Western blot (immunoblot) analysis is used to detect and quantify the presence of selected protein in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the antibodies that specifically bind

10   the immunogen protein. The anti-immunogen antibodies specifically bind to the protein on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (e.g., labeled sheep anti-mouse antibodies) that specifically bind to the anti-immunogen antibodies.

        Other assay formats include liposome immunoassays (LIA), which use

15   liposomes designed to bind specific molecules (e.g., antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (see Monroe et al., Amer. Clin. Prod. Rev. 5:34-41 (1986)).

Reduction of non-specific binding

20   One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate

25   with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

Labels

30   The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be

applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (e.g., DYNABEADS™), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine,

5   and the like), radiolabels (e.g., ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic beads (e.g., polystyrene, polypropylene, latex, etc.).

The label may be coupled directly or indirectly to the desired component of

10   the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

Non-radioactive labels are often attached by indirect means. Generally, a

15   ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to another molecules (e.g., streptavidin) molecule, which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable combination with antibodies that recognize a selected protein, or secondary antibodies that

20   recognize a primary antibody.

The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its

25   derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labeling or signal producing systems that may be used, see U.S. Patent No. 4,391,904.

Means of detecting labels are well known to those of skill in the art. Thus, for

30   example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, by means of photographic film, by the use of electronic detectors such as charge coupled devices (CCDs)

or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while

5   various conjugated beads appear the color of the bead.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the

10   target antibody is detected by simple visual inspection.

CELLULAR TRANSFECTION AND GENE THERAPY

The present invention provides the nucleic acids of a selected protein for the transfection of cells *in vitro* and *in vivo*. These nucleic acids can be inserted into any of a

15   number of well-known vectors for the transfection of target cells and organisms as described below. The nucleic acids are transfected into cells, *ex vivo* or *in vivo*, through the interaction of the vector and the target cell. The nucleic acid, under the control of a promoter, then expresses a protein of the present invention, thereby mitigating the effects of absent, partial inactivation, or abnormal expression of a gene, particularly as it relates to T cell activation

20   and migration. The compositions are administered to a patient in an amount sufficient to elicit a therapeutic response in the patient. An amount adequate to accomplish this is defined as "therapeutically effective dose or amount."

Such gene therapy procedures have been used to correct acquired and inherited genetic defects, cancer, and other diseases in a number of contexts. The ability to

25   express artificial genes in humans facilitates the prevention and/or cure of many important human diseases, including many diseases which are not amenable to treatment by other therapies (for a review of gene therapy procedures, *see* Anderson, *Science* 256:808-813 (1992); Nabel & Felgner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Mulligan, *Science* 926-932 (1993); Dillon, TIBTECH 11:167-175

30   (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1998); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada *et al.,* in *Current Topics in Microbiology and Immunology* (Doerfler & Böhm eds., 1995); and Yu *et al., Gene Therapy* 1:13-26 (1994)).

PHARMACEUTICAL COMPOSITIONS AND ADMINISTRATION

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered (e.g., nucleic acid, protein, modulatory compounds or

5   transduced cell), as well as by the particular method used to administer the composition. Accordingly, there are a wide variety of suitable formulations of pharmaceutical compositions of the present invention (*see, e.g., Remington's Pharmaceutical Sciences*, 17th ed., 1989). Administration can be in any convenient manner, e.g., by injection, oral administration, inhalation, transdermal application, or rectal administration.

10   Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the packaged nucleic acid suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include

15   one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, microcrystalline cellulose, gelatin, colloidal silicon dioxide, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a

20   flavor, e.g., sucrose, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art.

The compound of choice, alone or in combination with other suitable components, can be made into aerosol formulations (i.e., they can be "nebulized") to be

25   administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions,

30   which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or

intrathecally. Parenteral administration and intravenous administration are the preferred methods of administration. The formulations of commends can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials.

Injection solutions and suspensions can be prepared from sterile powders,
5   granules, and tablets of the kind previously described. Cells transduced by nucleic acids for *ex vivo* therapy can also be administered intravenously or parenterally as described above.

The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial therapeutic response in the patient over time. The dose will be determined by the efficacy of the particular vector employed and the condition
10   of the patient, as well as the body weight or surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular vector, or transduced cell type in a particular patient.

In determining the effective amount of the vector to be administered in the
15   treatment or prophylaxis of conditions owing to diminished or aberrant expression of the selected protein, the physician evaluates circulating plasma levels of the vector, vector toxicities, progression of the disease, and the production of anti-vector antibodies. In general, the dose equivalent of a naked nucleic acid from a vector is from about 1 μg to 100 μg for a typical 70 kilogram patient, and doses of vectors which include a retroviral particle are
20   calculated to yield an equivalent amount of therapeutic nucleic acid.

For administration, compounds and transduced cells of the present invention can be administered at a rate determined by the LD-50 of the inhibitor, vector, or transduced cell type, and the side-effects of the inhibitor, vector or cell type at various concentrations, as applied to the mass and overall health of the patient. Administration can be accomplished via
25   single or divided doses.

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

30

Example 1: Identification of EDG1 and other genes involved in modulation of T cell activation and migration

     A.   *Introduction*

          In this study, an approach to identify new targets for immune suppressive

5   drugs is provided. It is known that following T cell activation, expression of numerous cell surface markers such as CD25, CD69, and CD40L are upregulated. CD69 has been shown to be an early activation marker in T, B, and NK cells. CD69 is a disulfide-linked dimer. It is not expressed in resting lymphocytes but appears on T, B and NK cells after activation in vitro. Its relevance as a TCR signaling outcome has been validated using T cell deficient in

10   certain key signaling molecules such as LAT and SLP76 (Yablonski, *supra*). Furthermore, re-introducing SLP76 to the deficient cells results in restoration of CD69 expression. CD69 upregulation was therefore to be used to monitor TCR signal transduction. The rationale of the functional genomics screen was then to identify cell clones whose CD69 upregulation was repressed following introduction of a retroviral cDNA library. The library members

15   conferring such repression would then represent immune modulators that function to block TCR signal transduction.

     b.   *Results*

          Several T cell lines, including Jurkat, HPB-ALL, HSB-2 and PEER were

20   tested for the presence of surface CD3, CD25, CD28, CD40L, CD69, CD95, and CD95L. Those that express CD3 were cultured with anti-CD3 or anti-TCR to crosslink the TCR and examined for the upregulation of CD69. Jurkat T cell line was selected for its ability to upregulate CD69 in response to crosslinking of their TCR with a kinetics mimicking that of primary T lymphocytes (data not shown). The population of Jurkat cells was sorted for low

25   basal and highly inducible CD69 expression following anti-TCR stimulation. Clone 4D9 was selected because CD69 in this clone was uniformly and strongly induced following TCR stimulation in 24 hours.

          In order to regulate the expression of the retroviral library, the Tet-Off system was used. Basically, cDNA inserts in the retroviral library were cloned behind the

30   tetracycline regulatory element (TRE) and the minimal promoter of TK. Transcription of the cDNA inserts were then dependent on the presence of tetracycline-controlled trans-activator (tTA), a fusion of Tet repression protein and the VP16 activation domain, and the absence of tetracyaline or its derivatives such as doxycycline (Dox). To shut off the cDNA expression, one can simply add doxycycline in the medium. To obtain a Jurkat clone stably expresses

tTA, retroviral LTR-driven tTA was introduced in conjunction with a TRE-dependent reporter construct, namely TRA-Lyt2. Through sorting of Lyt2 positive cells in the absence of Dox and Lyt2 negative cells in the presence of Dox , coupled with clonal evaluation, a derivative of Jurkat clone 4D9 was obtained, called 4D9#32, that showed the best Dox

5   regulation of Lyt2 expression.

Positive controls: ZAP70 is a positive regulator of T cell activation. A kinase-inactivated (KI) ZAP70 and a truncated ZAP70 (SH2 N+ C) were subcloned into the retroviral vector under TRE control. ZAP70 SH2 (N+ C) and ZAP70 KI both inhibited TCR-

10   induced CD69 expression. Consistent with the published report on dominant negative forms of ZAP70 on NFAT activity, the truncated protein is also a more potent inhibitor of CD69 induction. In addition, the higher protein expression, as shown by adjusting GFP-gating, the stronger the inhibition was. When one puts the marker M1 at bottom 1% of the uninfected cells, one has a 40% likelihood of obtaining cells whose phenotype resembled that of ZAP70

15   SH2 (N+C). This translates into a 40:1 enrichment of the desired phenotype.

The CD69 inhibitory phenotype is dependent on expression of dominant negative forms of ZAP70. When Dox was added for 7 days before TCR was stimulated, there was no inhibition of CD69 expression. Analysis of cellular phenotype by FACS of GFP, which was produced from the bi-cistronic mRNA ZAP70 SH2 (N+C)-IRES-GFP, revealed a

20   lack of GFP+ cells. The lack of ZAP70 SH2 (N+C) expression in the presence of Dox was confirmed by Western.

Screening for cells lacking CD69 upregulation: Jurkat 4D9#32 cells were infected with cDNA libraries made form primary human lymphoid organs such as thymus,

25   spleen, lymph node and bone marrow. The library complexity was 5×10^7 and was built on the TRE vector. A total of 7.1×10^8 cells were screened with an infection rate of 52%, as judged by parallel infection of the same cells with TRA-dsGFP (data not shown). After infection, the cells will be stimulated with the anti-TCR antibody C305 for overnight and sorted for CD69 low and CD3+ phenotype by FACS. If the sorting gate was set to include

30   the bottom 3% cells based on the single parameter of CD69 level, 2/3 cells in the sorting gate lacked TCR/CD3 complex, which explained their refractory to stimulation. The second parameter of CD3 expression was then incorporated. Even though there was a significant reduction of CD3/TCR complex on the surface following receptor-mediated internalization, the CD3- population was still distinguishable from the CD3+ population. The resulting sort

gate contained 1% of the total cells, which translated into a 100-fold enrichment based on cell

numbers. The recovered cells with CD69 low CD3+ phenotype were allowed to rest in

complete medium for 5 days before being stimulated again for a new round of sorting. In

subsequent round of sortings, the sort gate was always maintained to contain the equivalent

5   of 1% of the unsorted control population. Obvious enrichment was achieved after 3 rounds

of reiterative sorting. Cells with the desired phenotype increased from 1% to 22.3%. In

addition, the overall population's geometric mean for CD69 was also reduced.

In order to ascertain that the phenotype was due to expression of the cDNA

library rather than entirely due to spontaneous or retroviral insertion-mediated somatic

10   mutation, the cells recovered after the third round of sorting were split into two halves. One

half of the cells were grown in the absence of Dox while the other half in the presence of

Dox. A week later, CD69 expression was compared following anti-TCR stimulation. There

was a significant numbers of cells (11%) whose CD69 repression was lost in the presence of

Dox, suggesting that the CD69 inhibition phenotype was indeed caused by the expression of

15   library members. Single cell clones in conjunction with the fourth round of CD69 low CD3+

sorting (LLLL) were deposited.

In order to reduce the number of cells whose phenotype was not Dox-

regulatable, the half of the cells grown in the presence of Dox were subjected to a fourth

round of sorting for enrichment of CD69 high phenotype (LLLH). The cells recovered from

20   LLLH sort were cultured in the absence of Dox for subsequence sorting and single cell

cloning of CD69 low CD3+ phenotypes.

Dox regulation of CD69 expression was expressed as the ratio of geometric

mean fluorescent intensity (GMFI) in the presence of Dox over that in the absence of Dox. In

uninfected cells, Dox had limited effect on the induction of CD69 expression so that the ratio

25   of GMFI (+Dox)/GMFI (-Dox) remained to be 1.00+/-0.25. The 2x standard deviation was

therefore used as a cut-off criterion and clones with a ratio above 1.5 were regarded as Dox-

regulated clones.

RNA samples were prepared from clones with Dox-regulatable phenotypes.

Using primers specific for the vector sequence flanking the cDNA library insert, the cDNA

30   insert of selected clones were captured by RT-PCR. Most clones generated only on DNA

band, whereas a few clones generated two or more bands. Sequencing analysis revealed that

the additional bands were caused by double or multiple insertions.

Characterization of proteins involved in T cell activation: Known TCR

regulators such as Lck, ZAP70, PLCγ1 and Raf were obtained. In addition, the BCR

regulator SYK was also uncovered. EDG1, a GPCR not previously known to be involved in

B and T cell activation, was also identified using this assay (see Figures 14-32).

5   Lck is a non-receptor protein tyrosine kinase. Its role in T cell development

and activation has been widely documented. So far, dominant negative form of Lck has no

been reported. Our discovery that over expression of the kinase-truncated form of Lck

caused inhibition of CD69, similar to the phenotype of Jurkat somatic mutant lacking Lck,

suggests that kinase deletion of Lck could also work as a dominant negative form of Lck.

10   The two ZAP70 hits ended at aa 262 and 269, respectively. They both missed

the catalytic domain. The deletions are very close to the positive control for the screen,

ZAP70 SH2 (N+C), which ended at aa 276. Since ZAP70 SH2 (N+C) was shown to be a

dominant negative protein, it appears that the two ZAP70 hits also behaved as dominant

negative proteins of ZAP70.

15   SYK is a non-receptor tyrosine kinase belonging to the SYK/ZAP70 family of

kinases. Since it has also been shown that the lack of SYK expression in Jurkat cells did not

appear to significantly alter the TCR-mediated responses compared with Jurkat clones

expressing SYK, it appears that the SYK hit obtained from our screen worked mainly to

block ZAP70 function. SYK's similarity to ZAP70 and its ability to associate with

20 phosphorylated TCR zeta chains also support this notion.

         PLCγ1 plays a crucial role in coupling T cell receptor ligation to IL-2 gene

expression in activated T lymphocytes. TCR engagement leads to rapid tyrosine

phosphorylation and activation of PLCγ1. The activated enzyme converts

phosphatidylinositol-4,5-bisphosphate (PIP2) to inositol-1,3,5-trisphosphate ((IP3) and

25 diacylglycerol (DAG). IP3 triggers intracellular Ca2+ increase and DAG is a potent activator

of protein kinase C (PKC). PLCγ1 has a split catalytic domain comprised of conserved X and

Y subdomains. Single point mutation in the catalytic X box completely abolished the

enzyme activity and also blocked IL-2 reporter gene expression when introduced into PLCγ-

deficient Jurkat cells. Our hit contained the PH domain and the N and C terminal SH2

30 domains of PLCγ1. Significantly this hit also deleted the crucial tyrosine Y783 between the

SH2 and SH3 domains. It was reported that Y783 was essential for coupling of TCR

stimulation to IL-2 promoter activation and that mutation of Y783 to F (phenylalanine)

generated a very potent dominant negative form of PLCγ1. Indeed, the original clone

encoding the PLCγ1 hit had the highest Dox +/- ratio for CD69 expression among all clones from the cDNA screen, indicating the strong repression of CD69 induction by the hit as well as the total de-repression in the absence of the hit. When introduced to naïve Jurkat cells, this fragment caused severe block of TCR-induced CD69 expression.

5   Raf is a MAP kinase kinase kinase. It interacts with Ras and leads to activation of the MAP kinase pathway. The Raf hit obtained also had a truncation of the kinase domain, creating a dominant negative form of the kinase. Other signaling molecules known to involve in TCR pathway were also discovered in our screen. They included PAG, CSK, SHP-1 and nucleolin.

10

Function in primary T lymphocytes: The relevance of the CD69 screen hits to physiological function of T cells was investigated in primary T lymphocytes. The hit was subcloned into a retroviral vector under a constitutively active promoter, followed by IRES-GFP. A protocol was also developed to couple successful retroviral infection to subsequence

15   T cell activation. Primary T lymphocytes are at the quiescent stage when isolated from healthy donors. In order to be infected by retrovirus, primary lymphocytes need to be activated to progress in cell cycle. Fresh peripheral blood lymphocytes (PBL) contained typically T cells and B cells. The combined CD4+ and CD8+ cells represented total T cell percentage, which was 81% in this particular donor. The remaining 19% CD4-CD8- cells

20   were B cells as stained by CD19 (data not shown). Upon culturing on anti-CD3 and anti-CD28 coated dishes, primary T lymphocytes were expanded and primary B cells and other cell types gradually died off in the culture. After infection, the culture contained virtually all T cells. Furthermore, primary T lymphocytes were successfully infected by retroviruses.

As seen with Jurkat cells (data not shown), GFP translated by way of IRES

25   was not as abundant as GFP translated using the conventional Kozak sequence (comparing GFP geometric mean from CRU5-IRES-GFP and CRU5-GFP). Nevertheless the percentage infection remained similar. Insertion of a gene in front of IRES-GFP further reduced the expression level of GFP, which was observed with cell lines (data not shown) and here primary T lymphocytes. After allowing cells to rest following infection, FACS sorted cells

30   were divided into two populations: GFP- and GFP+. The sorted cells were immediately put into culture. Anti-CD3 alone did not induce IL-2 production. This observation was consistent with previous report on freshly isolated primary T lymphocytes and confirmed the notion that prior culture and retroviral infection did not damage the physiological properties of these primary T lymphocytes. Addition of anti-CD28 in conjunction with anti-CD3 led to

robust IL-2 production with vector-infected cells and the GFP- population of LckDN and PLCγlDN-infected cells. The GFP+ cell population from LckDN and PLCγlDN-infected cells, however, were severely impaired in IL-2 production. As expect, the defect caused by LckDN and PLCγlDN can be completely rescued by stimulation using PMA and ionomycin.

5   Taken together, these results showed that Lck and PLCγl plays a role in IL-2 production from primary T lymphocytes, consistently with their involvement membrane proximal signaling events of T cell activation. These results also demonstrated a successful system to quickly validate hits from our functional genetic screens in primary cells.

10   **Use of CD69 upregulation in drug screening:** The discovery of important immune regulatory molecules from the B and T cell activation-induced CD69 upregulation validated the relevance of this cell-based assay. Essentially such a cell-based assay offers the opportunity to discover inhibitors of multiple targets such as Lck, ZAP70, PLCγl, and EDG family proteins such as EDG1. It is the equivalent of multiplexing enzymatic assays with the

15   additional advantage of cell permeability of compounds. It may even be possible to identify novel compounds that block adaptor protein functions. Towards this end, the FACS assay of cell surface CD69 expression was converted to a micro-titer plate based assay, for both T and B cell regulation assays.

20   In conclusion, the strategy presented in this study demonstrates a successful approach to discover and validate important immune regulators on a genome-wide scale. This approach, which requires no prior sequence information, provides a tool for functional cloning of regulators in numerous signal transduction pathways. For example, B cell activation-induced CD69 expression, IL-4-induced IgE class switch and TNF -induced NF-

25   kB reporter gene expression are all amendable to the genetic perturbation following introduction of retroviral cDNA libraries. The outlined strategy is less biased compared to forced introduction of a handful of signaling molecules discovered in other context such as growth factor signal transduction. It also opens the door for discovering peptide inhibitors of immune modulatory proteins by screening random peptide libraries, including cyclic

30   peptides, expressed from the retroviral vector.

C.   Methods

Cell culture: Human Jurkat T cells (clone N) were routinely cultured in RPMI 1640 medium supplemented with 10% fetal calf serum (Hyclone), penicillin and

streptamycin. Phoenix A cells were grown in DMEM supplemented with 10% fetal calf

serum, penicillin and streptamycin. To produce the tTA-Jurkat cell line, Jurkat cells were

infected with a retroviral construct which constitutively expresses the tetracycline

transactivator protein and a reporter construct which expresses LyT2 driven by a tetracycline

5   responsive element (TRE). The tTA-Jurkat cell population was optimized by sorting multiple

sounds for high TRE-dependent expression of LyT2 in the absence of Dox and strong

repression of LyT2 expression in the presence Dox. The cells were also sorted for maximal

anti-TCR induced expression of CD69. Doxycycline was used at a final concentration of

10ng/ml for at least 6 days to downregulate expression of cDNAs from the TRE promoter.

10

Transfection and infection: Phoenix A packaging cells were transfected with

retroviral vectors using calcium phosphate for 6 hours as standard protocols. After 24 hours,

supernatant was replaced with complete RPMI medium and virus was allowed to accumulate

for an additional 24 hours. Viral supernatant was collected, filtered through a 0.2μM filter

15   and mixed with Jurkat cells at a density of 2.5 x 10^5 cells/ml. Cells were spun at room

temperature for 3 hours at 3000 rpm, followed by overnight incubation at 37°C. Transfection

and infection efficiencies were monitored by GFP expression and functional analysis was

carried out 2-4 days after infection.

20   **Libraries:** RNA extracted from human lymph node, thymus, spleen and bone

marrow was used to produce two cDNA libraries; one random primed and directionally

cloned and the second non-directionally cloned and provided with 3 exogenous ATG in 3

frames. cDNAs were cloned into the pTRA-exs vector giving robust doxycycline–regulable

transcription of cDNAs from the TRE promoter. The total combined library complexity was 5

25   x 10^7 independent clones.

Stimulation: For CD69 upregulation experiments, tTA-Jurkat cells were split

to 2.5 x 10^5 cells/ml 24 hours prior to stimulation. Cells were spun and resuspended at 5 x

10^5 cells/ml in fresh complete RPMI medium in the presence of 100 ng/ml C305 (anti-Jurkat

30   clonotypic TCR) or 5 ng/ml PMA hybridoma supernatant for 20-26 hours at 37°C, and then

assayed for surface CD69 expression.

Cell surface marker analysis: Jurkat-N cells were stained with an APC-conjugated mouse monoclonal anti-human CD69 antibody (Caltag) at 4°C for 20 minutes and analyzed using a Facscalibur instrument (Becton Dickinson) with Cellquest software. Cell sorts were performed on a MoFlo (Cytomation).

5

cDNA screen: Phoenix A packaging cells were transfected with a mixture of the two tTA regulated retroviral pTRA-exs cDNA libraries. Supernatant containing packaged viral particles was used to infect tTA-Jurkat cells with an efficiency of ~85%. After 4 days of cDNA expression, library infected cells were stimulated with 0.3 μg/ml C305 for 20-26

10   hours, stained with APC-conjugated anti-CD69, and lowest CD69-expressing cells still expressing CD3 ($CD69^{low}CD3^+$) were isolated using a fluorescence activated cell sorter. Sorting was repeated over multiple rounds with a 6-day rest period between stimulations until the population was significantly enriched for non-responders. Single cells were deposited from 4 separate rounds of sorting. Cell clones were expanded in the presence and absence of

15   Dox, stimulated and analyzed for CD69 upregulation.

Isolation of cDNA inserts: PCR primers were designed to amplify cDNA inserts from both libraries and did not amplify Lyt2 that was also under TRE regulation. The primers used contained flanking BstXI sites for subsequent cloning to pTRA-IRES-GFP

20   vector. RT-PCR cloning was achieved with kits from Clontech or Life Technologies. The gel-purified RT-PCR products were submitted for sequencing directly and simultaneously digested for subcloning. Dominant negative ZAP70 (KI) and ZAP70SH2 (N+C) as well as selected hits from cDNA screens were subcloned to the retroviral pTRA-IRES-GFP vector. Selected hits form cDNA screens were also subcloned to CRU5-IRES-GFP for infection of

25   human primary T lymphocytes and examination of IL-2 production.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of

30   this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

SEQUENCE LISTING

A-raf-1

>gi|28820|emb|X04790.1|HSARAF1R Human mRNA for A-raf-1 oncogene (SEQ ID NO:1)
TGACCCAATAAGGGTGGAAGGCTGAGTCCCGCAGAGCCAATAACGAGAGTCCGAGAGGCGACGGAGGCGG
ACTCTGTGAGGAAACAAGAAGAGAGGCCCAAGATGGAGACGGCGGCGGCTGTAGCGGCGTGACAGGAGCC
CCATGGCACCTGCCCAGCCCCACCTCAGCCCATCTTGACAAAATCTAAGGCTCCATGGAGCCACCACGGG
GCCCCCCTGCCAATGGGGCCGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACCTGCCCAACAAGCA
ACGCACGGTGGTGACTGTCCGGGATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGAAGGTGCGG
GGTCTAAATCAGGACTGCTGTGTGGTCTACCGACTCATCAAGGGACGAAAGACGGTCACTGCCTGGGACA
CAGCCATTGCTCCCCTGGATGGCGAGGAGCTCATTGTCGAGGTCCTTGAAGATGTCCCGCTGACCATGCA
CAATTTTGTACGGAAGACCTTCTTCAGCCTGGCGTTCTGTGACTTCTGCCTTAAGTTTCTGTTCCATGGC
TTCCGTTGCCAAACCTGTGGCTACAAGTTCCACCAGCATTGTTCCTCCAAGGTCCCCACAGTCTGTGTTG
ACATGAGTACCAACCGCCAACAGTTCTACCACAGTGTCCAGGATTTGTCCGGAGGCTCCAGACAGCATGA
GGCTCCCTCGAACCGCCCCCTGAATGAGTTGCTAACCCCCCAGGGTCCCAGCCCCCGCACCCAGCACTGT
GACCCGGAGCACTTCCCCTTCCCTGCCCCAGCCAATGCCCCCCTACAGCGCATCCGCTCCACGTCCACTC
CCAACGTCCATATGGTCAGCACCACGGCCCCCATGGACTCCAACCTCATCCAGCTCACTGGCCAGAGTTT
CAGCACTGATGCTGCCGGTAGTAGAGGAGGTAGTGATGGAACCCCCGGGGGAGCCCCAGCCCAGCCAGC
GTGTCCTCGGGGAGGAAGTCCCCACATTCCAAGTCACCAGCAGAGCAGCGCGAGCGGAAGTCCTTGGCCG
ATGACAAGAAGAAAGTGAAGAACCTGGGGTACCGGGANTCAGGCTATTACTGGGAGGTACCACCCAGTGA
GGTGCAGCTGCTGAAGAGGATCGGGACGGGCTCGTTTGGCACCGTGTTTCGAGGGCGGTGGCATGGCGAT
GTGGCCGTGAAGGTGCTCAAGGTGTCCCAGCCCACAGCTGAGCAGGCCCAGGCTTTCAAGAATGAGATGC
AGGTGCTCAGGAAGACGCGACATGTCAACATCTTGCTGTTTATGGGCTTCATGACCCGGCCGGGATTTGC
CATCATCACACAGTGGTGTGAGGGCTCCAGCCTCTACCATCACCTGCATGTGGCCGACACACGCTTCGAC
ATGGTCCAGCTCATCGACGTGGCCCGGCAGACTGCCCAGGGCATGGACTACCTCCATGCCAAGAACATCA
TCCACCGAGATCTCAAGTCTAACAACATCTTCCTACATGAGGGGCTCACGGTGAAGATCGGTGACTTTGG
CTTGGCCACAGTGAAGACTCGATGGAGCGGGGCCCAGCCCTTGGAGCAGCCCTCAGGATCTGTGCTGTGG
ATGGCAGCTGAGGTGATCCGTATGCAGGACCCGAACCCCTACAGCTTCCAGTCAGACGTCTATGCCTACG
GGGTTGTGCTCTACGAGCTTATGACTGGCTCACTGCCTTACAGCCACATTGGCTGCCGTGACCAGATTAT
CTTTATGGTGGGCCGTGGCTATCTGTCCCCGGACCTCAGCAAAATCTCCAGCAACTGCCCCAAGGCCATG
CGGCGCCTGCTGTCTGACTGCCTCAAGTTCCAGCGGGAGGAGCGGCCCCTCTTCCCCCAGATCCTGGCCA
CAATTGAGCTGCTGCAACGGTCACTCCCCAAGATTGAGCGGAGTGCCTCGGAACCCTCCTTGCACCGCAC
CCAGGCCGATGAGTTGCCTGCCTGCCTACTCAGCGCAGCCCGCCTTGTGCCTTAGGCCCCGCCCAAGCCA
CCAGGGAGCCAATCTCAGCCCTCCACGCCAAGGAGCCTTGCCCACCAGCCAATCAATGTTCGTCTCTGCC
CTGATGCTGCCTCAGGATCCCCCATTCCCCACCCTGGGAGATGAGGGGGTCCCCATGTGCTTTTCCAGTT
CTTCTGGAATTGGGGGACCCCCGCCAAAGACTGAGCCCCCTGTCTCCTCCATCATTTGGTTTCCTCTTGG
CTTTGGGGATACTTCTAAATTTTGGGAGCTCCTCCATCTCCAATGGCTGGGATTTGTGGCAGGGATTCCA
CTCAGAACCTCTCTGGAATTTGTGCCTGATGTGCCTTCCACTGGATTTTGGGGTTCCCAGCACCCCATGT
GGATTTTGGGGGGTCCCTTTTGTGTCTCCCCCGCCATTCAAGGACTCCTCTCTTTCTTCACCAAGAAGCA
CAGAATTC

>gi|1340152|emb|CAA28476.1| ORF (A-raf) (AA 1-606) [Homo sapiens] (SEQ ID NO:2)
MEPPRGPPANGAEPSRAVGTVKVYLPNKQRTVVTVRDGMSVYDSLDKALKVRGLNQDCCVVYRLIKGRKT
VTAWDTAIAPLDGEELIVEVLEDVPLTMHNFVRKTFFSLAFCDFCLKFLFHGFRCQTCGYKFHQHCSSKV
PTVCVDMSTNRQQFYHSVQDLSGGSRQHEAPSNRPLNELLTPQGPSPRTQHCDPEHFPFPAPANAPLQRI
RSTSTPNVHMVSTTAPMDSNLIQLTGQSFSTDAAGSRGGSDGTPRGSPSPASVSSGRKSPHSKSPAEQRE
RKSLADDKKKVKNLGYRXSGYYWEVPPSEVQLLKRIGTGSFGTVFRGRWHGDVAVKVLKVSQPTAEQAQA
FKNEMQVLRKTRHVNILLFMGFMTRPGFAIITQWCEGSSLYHHLHVADTRFDMVQLIDVARQTAQGMDYL
HAKNIIHRDLKSNNIFLHEGLTVKIGDFGLATVKTRWSGAQPLEQPSGSVLWMAAEVIRMQDPNPYSFQS
DVYAYGVVLYELMTGSLPYSHIGCRDQIIFMVGRGYLSPDLSKISSNCPKAMRRLLSDCLKFQREERPLF
PQILATIELLQRSLPKIERSASEPSLHRTQADELPACLLSAARLVP

>gi|4502192|ref|NM_001654.1| Homo sapiens v-raf murine sarcoma 3611 viral oncogene homolog 1 (ARAF1), mRNA (SEQ ID NO:3)
ACGTGACCCTGACCCAATAAGGGTGGAAGGCTGAGTCCGCAGAGCCAATAACGAGAGTCCGAGAGGCGAC
GGAGGCGGACTCTGTGAGGAAACAAGAAGAGAGGCCCAAGATGGAGACGGCGGCGGCTGTAGCGGCGTGA
CAGGAGCCCCATGGCACCTGCCCAGCCCCACCTCAGCCCATCTTGACAAAATCTAAGGCTCCATGGAGCC
ACCACGGGGCCCCCCTGCCAATGGGGCCGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACCTGCCC

```
      AACAAGCAACGCACGGTGGTGACTGTCCGGGATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGA
      AGGTGCGGGGTCTAAATCAGGACTGCTGTGTGGTCTACCGACTCATCAAGGGACGAAAGACGGTCACTGC
      CTGGGACACAGCCATTGCTCCCCTGGATGGCGAGGAGCTCATTGTCGAGGTCCTTGAAGATGTCCCGCTG
      ACCATGCACAATTTTGTACGGAAGACCTTCTTCAGCCTGGCGTTCTGTGACTTCTGCCTTAAGTTTCTGT
  5   TCCATGGCTTCCGTTGCCAAACCTGTGGCTACAAGTTCCACCAGCATTGTTCCTCCAAGGTCCCCACAGT
      CTGTGTTGACATGAGTACCAACCGCCAACAGTTCTACCACAGTGTCCAGGATTTGTCCGGAGGCTCCAGA
      CAGCATGAGGCTCCCTCGAACCGCCCCCTGAATGAGTTGCTAACCCCCCAGGGTCCCAGCCCCCGCACCC
      AGCACTGTGACCCGGAGCACTTCCCCTTCCCTGCCCCAGCCAATGCCCCCCTACAGCGCATCCGCTCCAC
      GTCCACTCCCAACGTCCATATGGTCAGCACCACGGCCCCCATGGACTCCAACCTCATCCAGCTCACTGGC
 10   CAGAGTTTCAGCACTGATGCTGCCGGTAGTAGAGGAGGTAGTGATGGAACCCCCCGGGGGAGCCCCAGCC
      CAGCCAGCGTGTCCTCGGGGAGGAAGTCCCCACATTCCAAGTCACCAGCAGAGCAGCGCGAGCGGAAGTC
      CTTGGCCGATGACAAGAAGAAAGTGAAGAACCTGGGGTACCGGGACTCAGGCTATTACTGGGAGGTACCA
      CCCAGTGAGGTGCAGCTGCTGAAGAGGATCGGGACGGGCTCGTTTGGCACCGTGTTTCGAGGGCGGTGGC
      ATGGCGATGTGGCCGTGAAGGTGCTCAAGGTGTCCCAGCCCACAGCTGAGCAGGCCCAGGCTTTCAAGAA
 15   TGAGATGCAGGTGCTCAGGAAGACGCGACATGTCAACATCTTGCTGTTTATGGGCTTCATGACCCGGCCG
      GGATTTGCCATCATCACACAGTGGTGTGAGGGCTCCAGCCTCTACCATCACCTGCATGTGGCCGACACAC
      GCTTCGACATGGTCCAGCTCATCGACGTGGCCCGGCAGACTGCCCAGGGCATGGACTACCTCCATGCCAA
      GAACATCATCCACCGAGATCTCAAGTCTAACAACATCTTCCTACATGAGGGGCTCACGGTGAAGATCGGT
      GACTTTGGCTTGGCCACAGTGAAGACTCGATGGAGCGGGGCCCAGCCCTTGGAGCAGCCCTCAGGATCTG
 20   TGCTGTGGATGGCAGCTGAGGTGATCCGTATGCAGGACCCGAACCCCTACAGCTTCCAGTCAGACGTCTA
      TGCCTACGGGGTTGTGCTCTACGAGCTTATGACTGGCTCACTGCCTTACAGCCACATTGGCTGCCGTGAC
      CAGATTATCTTTATGGTGGGCCGTGGCTATCTGTCCCCGGACCTCAGCAAAATCTCCAGCAACTGCCCCA
      AGGCCATGCGGCGCCTGCTGTCTGACTGCCTCAAGTTCCAGCGGGAGGAGCGGCCCCTCTTCCCCCAGAT
      CCTGGCCACAATTGAGCTGCTGCAACGGTCACTCCCCAAGATTGAGCGGAGTGCCTCGGAACCCTCCTTG
 25   CACCGCACCCAGGCCGATGAGTTGCCTGCCTGCCTACTCAGCGCAGCCCGCCTTGTGCCTTAGGCCCCGC
      CCAAGCCACCAGGGAGCCAATCTCAGCCCTCCACGCCAAGGAGCCTTGCCCACCAGCCAATCAATGTTCG
      TCTCTGCCCTGATGCTGCCTCAGGATCCCCCATTCCCCACCCTGGGAGATGAGGGGGTCCCCATGTGCTT
      TTCCAGTTCTTCTGGAATTGGGGGGACCCCCGCCAAAGACTGAGCCCCCTGTCTCCTCCATCATTTGGTTT
      CCTCTTTGGCTTTGGGGATACTTCTAAATTTTGGGAGCTCCTCCATCTCCAATGGCTGGGATTTGTGGCA
 30   GGGATTCCACTCAGAACCTCTCTGGAATTTGTGCCTGATGTGCCTTCCACTGGATTTTGGGGTTCCCAGC
      ACCCCATGTGGATTTTGGGGGTCCCTTTTGTGTCTCCCCGCCATTCAAGGACTCCTCTCTTTCTTCACC
      AAGAAGCACAGAATTC
```

>gi|4502193|ref|NP_001645.1| v-raf murine sarcoma 3611 viral oncogene
homolog 1; Oncogene ARAF1 [Homo sapiens] (SEQ ID NO:4)

```
MEPPRGPPANGAEPSRAVGTVKVYLPNKQRTVVTVRDGMSVYDSLDKALKVRGLNQDCCVVYRLIKGRKT
VTAWDTAIAPLDGEELIVEVLEDVPLTMHNFVRKTFFSLAFCDFCLKFLFHGFRCQTCGYKFHQHCSSKV
PTVCVDMSTNRQQFYHSVQDLSGGSRQHEAPSNRPLNELLTPQGPSPRTQHCDPEHFPFPAPANAPLQRI
RSTSTPNVHMVSTTAPMDSNLIQLTGQSFSTDAAGSRGGSDGTPRGSPSPASVSSGRKSPHSKSPAEQRE
RKSLADDKKKVKNLGYRDSGYYWEVPPSEVQLLKRIGTGSFGTVFRGRWHGDVAVKVLKVSQPTAEQAQA
FKNEMQVLRKTRHVNILLFMGFMTRPGFAIITQWCEGSSLYHHLHVADTRFDMVQLIDVARQTAQGMDYL
HAKNIIHRDLKSNNIFLHEGLTVKIGDFGLATVKTRWSGAQPLEQPSGSVLWMAAEVIRMQDPNPYSFQS
DVYAYGVVLYELMTGSLPYSHIGCRDQIIFMVGRGYLSPDLSKISSNCPKAMRRLLSDCLKFQREERPLF
PQILATIELLQRSLPKIERSASEPSLHRTQADELPACLLSAARLVP
```

Lck

>gi|775207|gb|U23852.1|HSU23852 Human T-lymphocyte specific protein
tyrosine kinase p56lck (lck) aberrant mRNA, complete cds (SEQ ID NO:5)

```
GGTTAGGCCAGGAGGACCATGTGAATGGGGCCAGAGGGCTCCCGGGCTGGGCAGGGACCATGGGCTGTGG
CTGCAGCTCACACCCGGAAGATGACTGGATGGAAAACATCGATGTGTGTGAGAACTGCCATTATCCCATA
GTCCCACTGGATGGCAAGGGCACGCTGCTCATCCGAAATGGCTCTGAGGTGCGGGACCCACTGGTTACCT
ACGAAGGCTCCAATCCGCCGGCTTCCCCACTGCAAGACAACCTGGTTATCGCTCTGCACAGCTATGAGCC
CTCTCACGACGGAGATCTGGGCTTTGAGAAGGGGGAACAGCTCCGCATCCTGGAGCAGAGCGGCGAGTGG
TGGAAGGCGCAGTCCCTGACCACGGGCCAGGAAGGCTTCATCCCCTTCAATTTTGTGGCCAAAGCGAACA
GCCTGGAGCCCGAACCCTGGTTCTTCAAGAACCTGAGCCGCAAGGACGCGGAGCGGCAGCTCCTGGCGCC
CGGGAACACTCACGGCTCCTTCCTCATCCGGGAGAGCGAGAGCACCGCGGGATCGTTTTCACTGTCGGTC
CGGGACTTCGACCAGAACCAGGGAGAGGTGGTGAAACATTACAAGATCCGTAATCTGGACAACGGTGGCT
TCTACATCTCCCCTCGAATCACTTTTCCCGGCCTGCATGAACTGGTCCGCCATTACACCAATGCTTCAGA
TGGGCTGTGCACACGGTTGAGCCGCCCCTGCCAGACCCAGAAGCCCCAGAAGCCGTGGTGGGAGGACGAG
```

```
       TGGGAGGTTCCCAGGGAGACGCTGAAGCTGGTGGAGCGGCTGGGGGCTGGACAGTTCGGGGAGGTGTGGA
       TGGGGTACTACAACGGGCACACGAAGGTGGCGGTGAAGAGCCTGAAGCAGGGCAGCATGTCCCCCGACGC
       CTTCCTGGCCGAGGCCAACCTCATGAAGCAGCTGCAACACCAGCGGCTGGTTCGGCTCTACGCTGTGGTC
       ACCCAGGAGCCCATCTACATCATCACTGAATACATGGAGAATGGGAGTCTAGTGGATTTTCTCAAGACCC
5      CTTCAGGCATCAAGTTGACCATCAACAAACTCCTGGACATGGCAGCCCAAGTAAGGAGACTGGGGAGGGG
       GGCTGGGCAAGGGAACAGACCAGTGACGTGAAGACATCTGGCTCAGGACCGCTGATCTGTGTTTGGCCTG
       CAGATTGCAGAAGGCATGGCATTCATTGAAGAGCGGAATTATATTCATCGTGACCTTCGGGCTGCCAACA
       TTCTGGTGTCTGACACCCTGAGCTGCAAGATTGCAGACTTTGGCCTAGCACGCCTCATTGAGGACAACGA
       GTACACAGCCAGGGAGGGGGCCAAGTTTCCCATTAAGTGGACAGCGCCAGAAGCCATTAACTACGGGACA
10     TTCACCATCAAGTCAGATGTGTGGTCTTTTGGGATCCTGCTGACGGAAATTGTCACCCACGGCCGCATCC
       CTTACCCAGGGATGACCAACCCGGAGGTGATTCAGAACCTGGAGCGAGGCTACCGCATGGTGCGCCCTGA
       CAACTGTCCAGAGGAGCTGTACCAACTCATGAGGCTGTGCTGGAAGGAGCGCCCAGAGGACCGGCCCACC
       TTTGACTACCTGCGCAGTGTGCTGGAGGACTTCTTCACGGCCACAGAGGGCCAGTACCAGCCTCAGCCTT
       GAGAGGCCTTGAGAGGCCCTGGGGTTCTCCCCCTTTCTCTCCAGCCTGACTTGGGGAGATGGAGTTCTTG
15     TGCCATAGTCACATTGGCCATGCACATATGGACTCTGCACATGAATCCCACCCACATGTGACACATATGC
       ACCTTGTGTCTGTACACGTGTCCTGTAGTGTTGCGGACTCTGCACATGTCTTGTACATGTGTAGCCTGTG
       CATGTATGTCTTGGACACTGTACAAGGTACCCCTTTCTGGCTCTCCCATTTCCTGAGACCACAGAGAGAG
       GGGAGAAGCCTGGGATTGACAGAAGCTTCTGCCCACCTACTTTTCTTTCCTCAGATCATCCAGAAGTTCC
       TCAAGGGCCAGGACTTTATCTAATACCTCTGTGTGCTCCTCCTTGGTGCCTGGCCTGGCACACATCAGGA
20     GTTCAATAAATGTCTGTTGATGACTGCCG
```

>gi|775208|gb|AAC50287.1| p56lck (SEQ ID NO:6)
MGCGCSSHPEDDWMENIDVCENCHYPIVPLDGKGTLLIRNGSEVRDPLVTYEGSNPPASPLQDNLVIALH
SYEPSHDGDLGFEKGEQLRILEQSGEWWKAQSLTTGQEGFIPFNFVAKANSLEPEPWFFKNLSRKDAERQ
25   LLAPGNTHGSFLIRESESTAGSFSLSVRDFDQNQGEVVKHYKIRNLDNGGFYISPRITFPGLHELVRHYT
NASDGLCTRLSRPCQTQKPQKPWWEDEWEVPRETLKLVERLGAGQFGEVWMGYYNGHTKVAVKSLKQGSM
SPDAFLAEANLMKQLQHQRLVRLYAVVTQEPIYIITEYMENGSLVDFLKTPSGIKLTINKLLDMAAQVRR
LGRGAGQGNRPVT

30

ZAP70

>gi|340038|gb|L05148.1|HUMTYRKIN Human protein tyrosine kinase related mRNA
sequence (SEQ ID NO:7)
```
35     GGAATAGGTTAGTTTCAGACAAGCCTGCTTGCCGGAGCTCAGCAGACACCAGGCCTTCCGGGCAGGCCTG
       GCCCACCGTGGGCCTCAGAGCTGCTGCTGGGGCATTCAGAACCGGCTCTCCATTGGCATTGGGACCAGAG
       ACCCCGCAAGTGGCCTGTTTGCCTGGACATCCACCTGTACGTCCCCAGGTTTCGGGAGGCCCAGGGGCGA
       TGCCAGACCCCGCGGCGCACCTGCCCTTCTTCTACGGCAGCATCTCGCGTGCCGAGGCCGAGGAGCACCT
       GAAGCTGGCGGGCATGGCGGACGGGCTCTTCCTGCTGCGCCAGTGCCTGCGCTCGCTGGGCGGCTATGTG
40     CTGTCGCTCGTGCACGATGTGCGCTTCCACCACTTTCCCATCGAGCGCCAGCTCAACGGCACCTACGCCA
       TTGCCGGCGGCAAAGCGCACTGTGGACCGGCAGAGCTCTGCGAGTTCTACTCGCGCGACCCCGACGGGCT
       GCCCTGCAACCTGCGCAAGCCGTGCAACCGGCCGTCGGGCCTCGAGCCGCAGCCGGGGGTCTTCGACTGC
       CTGCGAGACGCCATGGTGCGTGACTACGTGCGCCAGACGTGGAAGCTGGAGGGCGAGGCCCTGGAGCAGG
       CCATCATCAGCCAGGCCCCGCAGGTGGAGAAGCTCATTGCTACGACGGCCCACGAGCGGATGCCCTGGTA
45     CCACAGCAGCCTGACGCGTGAGGAGGCCGAGCGCAAACTTTACTCTGGGGCGCAGACCGACGGCAAGTTC
       CTGCTGAGGCCGCGGAAGGAGCAGGGCACATACGCCCTGTCCCTCATCTATGGGAAGACGGTGTACCACT
       ACCTCATCAGCCAAGACAAGGCGGGCAAGTACTGCATTCCCGAGGGCACCAAGTTTGACACGCTCTGGCA
       GCTGGTGGAGTATCTGAAGCTGAAGGCGGACGGGCTCATCTACTGCCTGAAGGAGGCCTGCCCCAACAGC
       AGTGCCAGCAACGCCTCAGGGGCTGCTGCTCCCACACTCCAGCCCACCATCCACGTTGACTCATCCTC
50     AGAGACGAATCGACACCCTCAACTCAGATGGATACACCCCTGAGCCAGCACGCATAACGTCCCCAGACAA
       ACCGCGGCCGATGCCCATGGACACGAGCGTGTATGAGAGCCCCTACAGCGACCCAGAGGAGCTCAAGGAC
       AAGAAGCTCTTCCTGAAGCGCGATAACCTCCTCATAGCTGACATTGAACTTGGCTGCGGCAACTTTGGCT
       CAGTGCGCCAGGGCGTGTACCGCATGCGCAAGAAGCAGATCGACGTGGCCATCAAGGTGCTGAAGCAGGG
       CACGGAGAAGGCAGACACGGAAGAGATGATGCGCGAGGCGCAGATCATGCACCAGCTGGACAACCCCTAC
55     ATCGTGCGGCTCATTGGCGTCTGCCAGGCCGAGGCCCTCATGCTGGTCATGGAGATGGCTGGGGGCGGGC
       CGCTGCACAAGTTCCTGGTCGGCAAGAGGGAGGAGATCCCTGTGAGCAATGTGGCCGAGCTGCTGCACCA
       GGTGTCCATGGGGATGAAGTACCTGGAGGAGAAGAACTTTGTGCACCGTGACCTGGCGGCCCGCAACGTC
       CTGCTGGTTAACCGGCACTACGCCAAGATCAGCGACTTTGGCCTCTCCAAAGCACTGGGTGCCGACGACA
       GCTACTACACTGCCCGCTCAGCAGGGAAGTGGCCGCTCAAGTGGTACGCACCCGAATGCATCAACTTCCG
60     CAAGTTCTCCAGCCGCAGCGATGTCTGGAGCTATGGGGTCACCATGTGGGAGGCCTTGTCCTACGGCCAG
       AAGCCCTACAAGAAGATGAAAGGGCCGGAGGTCATGGCCTTCATCGAGCAGGGCAAGCGGATGGAGTGCC
```

```
CACCAGAGTGTCCACCCGAACTGTACGCACTCATGAGTGACTGCTGGATCTACAAGTGGGAGGATCGCCC
CGACTTCCTGACCGTGGAGCAGCGCATGCGAGCCTGTTACTACAGCCTGGCCAGCAAGGTGGAAGGGCCC
CCAGGCAGCACACAGAAGGCTGAGGCTGCCTGTGCCTGAGCTCCCGCTGCCCAGGGGAGCCCTCCACGCC
GGCTCTTCCCCACCCTCAGCCCCACCCCAGGTCCTGCAGTCTGGCTGAGCCCTGCTTGGTTGTCTCCACA
CACAGCTGGGCTGTGGTAGGGGGTGTCTCAGGCCACACCGGCCTTGCATTGCCTGCCTGGCCCCCTGTCC
TCTCTGGCTGGGGAGCAGGGAGGTCCGGGAGGGTGCGGCTGTGCAGCCTGTCCTGGGCTGGTGGCTCCCG
GAGGGCCCTGAGCTGAGGGCATTGCTTACACGGATGCCTTCCCCTGGGCCCTGACATTGGAGCCTGGGCA
TCCTCAGGTGGTCAGGCGTAGATCACCAGAATAAACCCAGCTTCCCTCTTGAAAAAAAAAAAAAAAAAAA
AACC
```

>gi|18600045|ref|XP_047776.3| similar to Tyrosine-protein kinase ZAP-70 (70
kDa zeta-associated protein) (Syk-related tyrosine kinase) [Homo sapiens]
(SEQ ID NO:8)

```
MPDPAAHLPFFYGSISRAEAEEHLKLAGMADGLFLLRQCLRSLGGYVLSLVHDVRFHHFPIERQLNGTYA
IAGGKAHCGPAELCEFYSRDPDGLPCNLRKPCNRPSGLEPQPGVFDCLRDAMVRDYVRQTWKLEGEALEQ
AIISQAPQVEKLIATTAHERMPWYHSSLTREEAERKLYSGAQTDGKFLLRPRKEQGTYALSLIYGKTVYH
YLISQDKAGKYCIPEGTKFDTLWQLVEYLKLKADGLIYCLKEACPNSSASNASGAAAPTLPAHPSTLTHP
QRRIDTLNSDGYTPEPARITSPDKPRPMPMDTSVYESPYSDPEELKDKKLFLKRDNLLIADIELGCGNFG
SVRQGVYRMRKKQIDVAIKVLKQGTEKADTEEMMREAQIMHQLDNPYIVRLIGVCQAEALMLVMEMAGGG
PLHKFLVGKREEIPVSNVAELLHQVSMGMKYLEEKNFVHRDLAARNVLLVNRHYAKISDFGLSKALGADD
SYYTARSAGKWPLKWYAPECINFRKFSSRSDVWSYGVTMWEALSYGQKPYKKMKGPEVMAFIEQGKRMEC
PPECPPELYALMSDCWIYKWEDRPDFLTVEQRMRACYYSLASKVEGPPGSTQKAEAACA
```

SYK

>gi|479012|gb|L28824.1|HUMPTK Homo sapiens protein tyrosine kinase (Syk)
mRNA, complete cds (SEQ ID NO:9)

```
CTCTAAAGGCCGCGGGCCGGCGGCTGAGGCCACCCCGGCGGCGGCTGGAGAGCGAGGAGGAGCGGGTGGC
CCCGCGCTGCGCCCGCCCTCGCCTCACCTGGCGCAGGTGGACACCTGCCGAGGTGTGTGCCCTCCGGCCC
CTGAAGCATGGCCAGCAGCGGCATGGCTGACAGCGCCAACCACCTGCCCTTCTTTTTCGGCAACATCACC
CGGGAGGAGGCAGAAGATTACCTGGTCCAGGGGGGCATGAGTGATGGGCTTTATTTGCTGCGCCAGAGCC
GCAACTACCTGGGTGGCTTCGCCCTGTCCGTGGCCCACGGGAGGAAGGCACACCACTACACCATCGAGCG
GGAGCTGAATGGCACCTACGCCATCGCCGGTGGCAGGACCCATGCCAGCCCGCCGACCTCTGCCACTAC
CACTCCCAGGAGTCTGATGGCCTGGTCTGCCTCCTCAAGAAGCCCTTCAACCGGCCCCAAGGGGTGCAGC
CCAAGACTGGGCCCTTTGAGGATTTGAAGGAAAACCTCATCAGGGAATATGTGAAGCAGACATGGAACCT
GCAGGGTCAGGCTCTGGAGCAGGCCATCATCAGTCAGAAGCCTCAGCTGGAGAAGCTGATCGCTACCACA
GCCCATGAAAAAATGCCTTGGTTCCATGGAAAAATCTCTCGGGAAGAATCTGAGCAAATTGTCCTGATAG
GATCAAAGACAAATGGAAAGTTCCTGATCCGAGCCAGAGACAACAACGGCTCCTACGCCCTGTGCCTGCT
GCACGAAGGGAAGGTGCTGCACTATCGCATCGACAAAGACAAGACAGGGAAGCTCTCCATCCCCGAGGGA
AAGAAGTTCGACACGCTCTGGCAGCTAGTCGAGCATTATTCTTATAAAGCAGATGGTTTGTTAAGAGTTC
TTACTGTCCCATGTCAAAAAATCGGCACACAGGGAAATGTTAATTTTGGAGGCCGTCCACAACTTCCAGG
TTCCCATCCTGCGACTTGGTCAGCGGGTGGAATAATCTCAAGAATCAAATCATACTCCTTCCCAAAGCCT
GGCCACAGAAAGTCCTCCCCTGCCCAAGGGAACCGGCAAGAGTACTGTGTCATTCAATCCGTATGAGC
CAGAACTTGCACCCTGGGCTGCAGACAAAGGCCCCAGAGAGAAGCCCTACCCATGGACACAGAGGTGTA
CGAGAGCCCCTACGCGGACCCCGAGGAGATCAGGCCCAAGGAGGTTTACCTGGACCGAAAGCTGCTGACG
CTGGAAGACAAAGAACTGGGCTCTGGTAATTTTGGAACTGTGAAAAAGGGCTACTACCAAATGAAAAAAG
TTGTGAAAACCGTGGCTGTGAAAATACTGAAAAACGAGGCCAATGACCCCGCTCTTAAAGATGAGTTATT
AGCAGAAGCAAATGTCATGCAGCAGCTGGACAACCCGTACATCGTGCGGATGATCGGGATATGCGAGGCC
GAGTCCTGGATGCTGGTTATGGAGATGGCAGAACTTGGTCCCCTCAATAAGTATTTGCAGCAGAACAGAC
ATGTCAAGGATAAGAACATCATAGAACTGGTTCATCAGGTTTCCATGGGCATGAAGTACTTGGAGGAGAG
CAATTTTGTGCACAGAGATCTGGCTGCAAGAAATGTGTTGCTAGTTACCCAACATTACGCCAAGATCAGT
GATTTCGGACTTTCCAAAGCACTGCGTGCTGATGAAACTACTACAAGGCCCAGACCCATGGAAAGTGGC
CTGTCAAGTGGTACGCTCCGGAATGCATCAACTACTACAAGTTCTCCAGCAAAAGCGATGTCTGGAGCTT
TGGAGTGTTGATGTGGGAAGCATTCTCCTATGGGCAGAAGCCATATCGAGGGATGAAAGGAAGTGAAGTC
ACCGCTATGTTAGAGAAAGGAGAGCGGATGGGGTGCCCTGCAGGGTGTCCAAGAGAGATGTACGATCTCA
TGAATCTGTGCTGGACATACGATGTGGAAAACAGGCCCGGATTCGCAGCAGTGGAACTGCGGCTGCGCAA
TTACTACTATGACGTGGTGAACTAACCGCTCCCGCACCTGTCGGTGGCTGCCTTTGATCACAGGAGCAAT
CACAGGAAAATGTATCCAGAGGAATTGATTGTCAGCCACCTCCCTCTGCCAGTCGGGAGAGCCAGGCTTG
GATGAACATGCCCACAACTTGTCACCCAAAGCCTGTCCCAGGACTCACCCTCCACAAAGCAAAGGCAGT
CCCGGGAGAAAAGACGGATGGCAGGATCCAAGGGGCTAGCTGGATTTGTTTGTTTTCTTGTCTGTGTGAT
```

TTTCATACAGGTTATTTTTACGATCTGTTTCCAAATCCCTTTCATGTCTTTCCACTTCTCTGGGTCCCGG
GGTGCATTTGTTACTCATCGGGCCCAGGGACATTGCAGAGTGGCCTAGAGCACTCTCACCCCAAGCGGCC
TTTTCCAAATGCCCAAGGATGCCTTAGCATGTGACTCCTGAAGGAAGGCAAAGGCAGAGGAATTTGGCTG
CTTCTACGGCCATGAGACTGATCCCTGGCCACTGAAAAGCTTTCCTGACAATAAAAATGTTTTGAGGCTT
5   TAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTTTAGAGCACA

>gi|479013|gb|AAA36526.1| protein tyrosine kinase (SEQ ID NO:10)
MASSGMADSANHLPFFFGNITREEAEDYLVQGGMSDGLYLLRQSRNYLGGFALSVAHGRKAHHYTIEREL
NGTYAIAGGRTHASPADLCHYHSQESDGLVCLLKKPFNRPQGVQPKTGPFEDLKENLIREYVKQTWNLQG
10 QALEQAIISQKPQLEKLIATTAHEKMPWFHGKISREESEQIVLIGSKTNGKFLIRARDNNGSYALCLLHE
GKVLHYRIDKDKTGKLSIPEGKKFDTLWQLVEHYSYKADGLLRVLTVPCQKIGTQGNVNFGGRPQLPGSH
PATWSAGGIISRIKSYSFPKPGHRKSSPAQGNRQESTVSFNPYEPELAPWAADKGPQREALPMDTEVYES
PYADPEEIRPKEVYLDRKLLTLEDKELGSGNFGTVKKGYYQMKKVVKTVAVKILKNEANDPALKDELLAE
ANVMQQLDNPYIVRMIGICEAESWMLVMEMAELGPLNKYLQQNRHVKDKNIIELVHQVSMGMKYLEESNF
15 VHRDLAARNVLLVTQHYAKISDFGLSKALRADENYYKAQTHGKWPVKWYAPECINYYKFSSKSDVWSFGV
LMWEAFSYGQKPYRGMKGSEVTAMLEKGERMGCPAGCPREMYDLMNLCWTYDVENRPGFAAVELRLRNYY
YDVVN

20 **PLCG1**

>gi|4505868|ref|NM_002660.1| Homo sapiens phospholipase C, gamma 1
(formerly subtype 148) (PLCG1), mRNA (SEQ ID NO:11)
GGGGTGCCGCCGCCGCCGTTGCGCTTGCTCCCGGGCGGTCCTGGCCTGTGCCGCCGCCGCCCCCAGCGTC
25 GGAGCCATGGCGGGCGCCGCGTCCCCTTGCGCCAACGGCTGCGGGCCCGGCGCGCCCTCGGACGCCGAGG
TGCTGCACCTCTGCCGCAGCCTCGAGGTGGGCACCGTCATGACTTTGTTCTACTCCAAGAAGTCGCAGCG
ACCCGAGCGGAAGACCTTCCAGGTCAAGCTGGAGACGCGCCAGATCACGTGGAGCCGGGGCGCCGACAAG
ATCGAGGGGGCCATTGACATTCGTGAAATTAAGGAGATCCGCCCAGGGAAGACCTCACGGGACTTTGATC
GCTATCAAGAGGACCCAGCTTTCCGGCCGGACCAGTCACATTGCTTTGTCATTCTCTATGGAATGGAATT
30 TCGCCTGAAAACGCTGAGCCTGCAAGCCACATCTGAGGATGAAGTGAACATGTGGATCAAGGGCTTAACT
TGGCTGATGGAGGATACATTGCAGGCACCCACACCCTGCAGATTGAGAGGTGGCTCCGGAAGCAGTTTT
ACTCAGTGGATCGGAATCGTGAGGATCGTATATCAGCCAAGGACCTGAAGAACATGCTGTCCCAGGTCAA
CTACCGGGTCCCCAACATGCGCTTCCTCCGAGAGCGGCTGACGGACCTGGAGCAGCGCAGCGGGGACATC
ACCTACGGGCAGTTTGCTCAGCTGTACCGCAGCCTCATGTACAGCGCCCAGAAGACGATGGACCTCCCCT
35 TCTTGGAAGCCAGTACTCTGAGGGCTGGGGAGCGGCCGGAGCTTTGCCGAGTGTCCCTTCCTGAGTTCCA
GCAGTTCCTTCTTGACTACCAGGGGGAGCTGTGGGCTGTTGATCGCCTCCAGGTGCAGGAGTTCATGCTC
AGCTTCCTCCGAGACCCCCTTACGAGAGATCGAGGAGCCATACTTCTTCCTGGATGAGTTTGTCACCTTCC
TGTTCTCCAAAGAGAACAGTGTGTGGAACTCGCAGCTGGATGCAGTATGCCCGGACACCATGAACAACCC
TCTTTCCCACTACTGGATCTCCTCCTCGCACAACACGTACCTGACCGGGGACCAGTTCTCCAGTGAGTCC
40 TCCTTGGAAGCCTATGCTCGCTGCCTGCGGATGGGCTGTCGCTGCATTGAGTTGGACTGCTGGGACGGCC
CGGATGGGATGCCAGTTATTTACCATGGGCACACCCTTACCACCAAGATCAAGTTCTCAGATGTCCTGCA
CACCATCAAGGAGCATGCCTTTGTGGCCTCAGAGTACCCAGTCATCCTGTCCATTGAGGACCACTGCAGC
ATTGCCCAGCAGAGAAACATGGCCCAATACTTCAAGAAGGTGCTGGGGGACACACTCCTCACCAAGCCCG
TGGAGATCTCTGCCGACGGGCTCCCCTCACCCAACCAGCTTAAGAGGAAGATCCTCATCAAGCACAAGAA
45 GCTGGCTGAGGGCAGTGCCTACGAGGAGGTGCCTACATCCATGATGTACTCTGAGAACGACATCAGCAAC
TCTATCAAGAATGGCATCCTCTACCTGGAGGACCCTGTGAACCACGAATGGTATCCCCACTACTTTGTTC
TGACCAGCAGCAAGATCTACTACTCTGAGGAGACCAGCAGTGACCAGGGCAACGAGGATGAGGAGGAGCC
CAAGGAGGTCAGCAGCAGCACAGAGCTGCACTCCAATGAGAAGTGGTTCCATGGGAAGCTAGGGGCAGGG
CGTGACGGGCGTCACATCGCTGAGCGCCTGCTTACTGAGTACTGCATCGAGACCGGAGCCCCTGACGGCT
50 CCTTCCTCGTGCGAGAGAGTGAGACCTTCGTGGGCGACTACACGCTCTCTTTCTGGCGGAACGGGAAAGT
CCAGCACTGCCGTATCCACTCCCGGCAAGATGCTGGGACCCCCAAGTTCTTCTTGACAGACAACCTCGTC
TTTGACTCCCTCTATGACCTCATCACGCACTACCAGCAGGTGCCCCTGCGCTGTAATGAGTTTGAGATGC
GACTTTCAGAGCCTGTCCCACAGACCAACGCCCACGAGAGCAAAGAGTGGTACCACGCGAGCCTGACCAG
AGCACAGGCTGAGCACATGCTAATGCGCGTCCCTCGTGATGGGGCCTTCCTGGTGCGGAAGCGGAATGAA
55 CCCAACTCATATGCCATCTCTTTCCGGGCTGAGGGCAAGATCAAGCATTGCCGTGTCCAGCAAGAGGGCC
AGACAGTGATGCTAGGGAACTCGGAGTTCGACAGCCTTGTTGACCTCATCAGCTACTATGAGAAACACCC
GCTATACCGCAAGATGAAGCTGCGCTATCCCATCAACGAGGAGGCACTGGAGAAGATTGGCACAGCTGAG
CCTGACTACGGGGCCCTGTATGAGGGACGCAACCCTGGCTTCTATGTAGAGGCAAACCCTATGCCAACTT
TCAAGTGTGCAGTCAAAGCCCTCTTTGACTACAAGGCCCAGAGGGAGGACGAGCTGACCTTCATCAAGAG
60 CGCCATCATCCAGAATGTGGAGAAGCAAGAGGGAGGCTGGTGGCGAGGGGACTACGGAGGGAAGAAGCAG
CTGTGGTTCCCCATCAAACTACGTGGAAGAGATGGTCAACCCCGTGGCCCTGGAGCCGGAGAGGGAGCACT

TGGACGAGAACAGCCCCCTAGGGGACTTGCTGCGGGGGGTCTTGGATGTGCCGGCTTGTCAGATTGCCAT
CCGTCCTGAGGGCAAGAACAACCGGCTCTTCGTCTTCTCCATCAGCATGGCGTCGGTGGCCCACTGGTCC
CTGGATGTTGCTGCCGACTCACAGGAGGAGCTGCAGGACTGGGTGAAAAAGATCCGTGAAGTGGCCCAGA
CAGCAGACGCCAGGCTCACTGAAGGGAAGATAATGGAACGGAGGAAGAAGATTGCCCTGGAGCTCTCTGA
ACTTGTCGTCTACTGCCGGCCTGTTCCCTTTGATGAAGAGAAGATTGGCACAGAACGTGCTTGCTACCGG
GACATGTCATCCTTCCCGGAAACCAAGGCTGAGAAATACGTGAACAAGGCCAAAGGCAAGAAGTTCCTTC
AGTACAATCGACTGCAGCTCTCCCGCATCTACCCCAAGGGCCAGCGACTGGATTCCTCCAACTACGATCC
TTTGCCCATGTGGATCTGTGGCAGTCAGCTTGTGGCCCTCAACTTCCAGACCCCTGACAAGCCTATGCAG
ATGAACCAGGCCCTCTTCATGACGGGCAGGCACTGTGGCTACGTGCTGCAGCCAAGCACCATGCGGGATG
AGGCCTTCGACCCCTTTGACAAGAGCAGCCTCCGCGGGCTGGAGCCATGTGCCATCTCTATTGAGGTGCT
GGGGGCCCGACATCTGCCAAAGAATGGCCGAGGCATTGTGTGTCCTTTTGTGGAGATTGAGGTGGCTGGA
GCTGAGTATGACAGCACCAAGCAGAAGACAGAGTTTGTGGTGGACAATGGACTCAACCCTGTATGGCCAG
CCAAGCCCTTCCACTTCCAGATCAGTAACCCTGAATTTGCCTTTCTGCGCTTCGTGGTGTATGAGGAAGA
CATGTTTAGTGACCAGAATTTCCTGGCTCAGGCTACTTTCCCAGTAAAAGGCCTGAAGACAGGATACAGA
GCAGTGCCTTTGAAGAACAACTACAGTGAGGACCTGGAGTTGGCCTCCCTGCTGATCAAGATTGACATTT
TCCCTGCCAAGGAGAATGGTGACCTCAGTCCCTTCAGTGGTACGTCCCTGCGGGAGCGGGGCTCAGATGC
CTCAGGCCAGCTGTTTCATGGCCGAGCCCGGGAAGGCTCCTTTGAATCCCGCTACCAGCAGCCGTTTGAG
GACTTCCGCATCTCCCAGGAGCATCTCGCAGACCATTTTGACAGTCGAGAACGAAGGGCCCCAAGAAGGA
CTCGGGTCAATGGAGACAACCGCCTCTAGTTGTACCCCAGCCTCGTTGGAGAGCAGCAGGTGCTGTGCGC
CTTGTAGAATGCCGCGAACTGGGTTCTTTGGAAGCAGCCCCCTGTGGCGGCCTTCCGGGTCTCGCAGCCT
GAAGCCTGGATTCCAGCAGTGAATGCTAGACAGAAACCAAGCCATTAATGAGATGTTATTACTGTTTTGG
GCCTCCATGCCCCAGCTCTGGATGAAGGCAAAAACTGTACTGTGTTTCGCATTAAGCACACACATCTGGC
CCTGACTTCTGGAGATGGATCCTTCCATCTTGTGGGGCCAGGACCATGGCCGAAGCCCCTTGGAGAGAGA
GGCTGCCTCAGCCAGTGGCACAGGAGACTCCAAGGAGCTACTGACATTCCTAAGAGTGGAGGAGGAGGAG
GAGCCTTGCTGGGCCAGGGAAACAAAGTTTACATTGTCCTGTAGCTTTAAAACCACAGCTGGGCAGGG

>gi|4505869|ref|NP_002651.1| phospholipase C, gamma 1 (formerly subtype
148) [Homo sapiens] (SEQ ID NO:12)
MAGAASPCANGCGPGAPSDAEVLHLCRSLEVGTVMTLFYSKKSQRPERKTFQVKLETRQITWSRGADKIE
GAIDIREIKEIRPGKTSRDFDRYQEDPAFRPDQSHCFVILYGMEFRLKTLSLQATSEDEVNMWIKGLTWL
MEDTLQAPTPLQIERWLRKQFYSVDRNREDRISAKDLKNMLSQVNYRVPNMRFLRERLTDLEQRSGDITY
GQFAQLYRSLMYSAQKTMDLPFLEASTLRAGERPELCRVSLPEFQQFLLDYQGELWAVDRLQVQEFMLSF
LRDPLREIEEPYFFLDEFVTFLFSKENSVWNSQLDAVCPDTMNNPLSHYWISSSHNTYLTGDQFSSESSL
EAYARCLRMGCRCIELDCWDGPDGMPVIYHGHTLTTKIKFSDVLHTIKEHAFVASEYPVILSIEDHCSIA
QQRNMAQYFKKVLGDTLLTKPVEISADGLPSPNQLKRKILIKHKKLAEGSAYEEVPTSMMYSENDISNSI
KNGILYLEDPVNHEWYPHYFVLTSSKIYYSEETSSDQGNEDEEEPKEVSSSTELHSNEKWFHGKLGAGRD
GRHIAERLLTEYCIETGAPDGSFLVRESETFVGDYTLSFWRNGKVQHCRIHSRQDAGTPKFFLTDNLVFD
SLYDLITHYQQVPLRCNEFEMRLSEPVPQTNAHESKEWYHASLTRAQAEHMLMRVPRDGAFLVRKRNEPN
SYAISFRAEGKIKHCRVQQEGQTVMLGNSEFDSLVDLISYYEKHPLYRKMKLRYPINEEALEKIGTAEPD
YGALYEGRNPGFYVEANPMPTFKCAVKALFDYKAQREDELTFIKSAIIQNVEKQEGGWWRGDYGGKKQLW
FPSNYVEEMVNPVALEPEREHLDENSPLGDLLRGVLDVPACQIAIRPEGKNNRLFVFSISMASVAHWSLD
VAADSQEELQDWVKKIREVAQTADARLTEGKIMERRKKIALELSELVVYCRPVPFDEEKIGTERACYRDM
SSFPETKAEKYVNKAKGKKFLQYNRLQLSRIYPKGQRLDSSNYDPLPMWICGSQLVALNFQTPDKPMQMN
QALFMTGRHCGYVLQPSTMRDEAFDPFDKSSLRGLEPCAISIEVLGARHLPKNGRGIVCPFVEIEVAGAE
YDSTKQKTEFVVDNGLNPVWPAKPFHFQISNPEFAFLRFVVYEEDMFSDQNFLAQATFPVKGLKTGYRAV
PLKNNYSEDLELASLLIKIDIFPAKENGDLSPFSGTSLRERGSDASGQLFHGRAREGSFESRYQQPFEDF
RISQEHLADHFDSRERRAPRRTRVNGDNRL

PAG

>gi|7682683|gb|AF240634.1|AF240634 Homo sapiens phosphoprotein associated
with GEMs (PAG) mRNA, complete cds (SEQ ID NO:13)
ATGGGGCCCGCGGGGAGCCTGCTGGGCAGCGGACAGATGCAGATCACCCTGTGGGGAAGTCTGGCTGCTG
TCGCCATTTTCTTCGTCATCACCTTCCTCATCTTCCTGTGCTCTAGTTGTGACAGGGAAAAGAAGCCGCG
ACAGCATAGTGGGGACCATGAGAACCTGATGAACGTGCCTTCAGACAAGGAGATGTTCAGCCGTTCAGTT
ACTAGCCTGGCAACAGATGCTCCTGCCAGCAGTGAGCAGAATGGGGCACTCACCAATGGGACATTCTTT
CAGAGGACAGTACTCTGACCTGCATGCAGCATTACGAGGAAGTCCAGACATCGGCCTCGGATCTGCTGGA
TTCCCAGGACAGCACAGGGAAACCAAAATGTCATCAGAGTCGGGAGCTGCCCAGAATCCCTCCCGAGAGC
GCAGTGGATACCATGCTCACGGCGAGAAGTGTGGACGGGGACCAGGGGCTGGGGATGGAAGGGCCCTATG
AAGTGCTCAAGGACAGCTCCTCCCAAGAAAACATGGTGGAGGACTGCTTGTATGAAACTGTGAAAGAGAT

CAAGGAGGTGGCTGCAGCTGCACACCTGGAGAAAGGCCACAGTGGCAAGGCAAAATCTACTTCTGCCTCG
AAAGAGCTCCCAGGGCCCCAGACTGAAGGCAAAGCTGAGTTTGCTGAATATGCCTCGGTGGACAGAAACA
AAAAATGTCGCCAAAGTGTTAATGTAGAGAGTATCCTTGGAAATTCATGTGATCCAGAAGAGGAGGCCCC
ACCACCTGTCCCTGTTAAGCTTCTGGACGAGAATGAAAACCTTCAGGAGAAGGAAGGGGGAGAGGCGGAA
5   GAGAGTGCCACAGACACGACCAGTGAAACTAACAAGAGATTTAGCTCATTGTCATACAAGTCTCGGGAAG
AAGACCCCACTCTCACAGAAGAAGAGATCTCAGCTATGTACTCATCAGTAAATAAACCTGGACAGTTAGT
GAATAAATCGGGGCAGTCGCTTACAGTTCCGGAGTCCACCTACACCTCCATTCAAGGGGACCCACAGAGG
TCACCCTCCTCCTGTAATGATCTCTATGCTACTGTTAAAGACTTCGAAAAAACTCCAAACAGCACACTTC
CACCAGCAGGGAGGCCCAGCGAGGAGCCAGAGCCTGATTATGAAGCGATACAGACTCTCAACAGAGAGGA
10   AGAAAAGGCCACCCTGGGGACCAATGGCCACCACGGTCTCGTCCCAAAGGAGAACGACTACGAGAGCATA
AGTGACTTGCAGCAAGGCAGAGATATTACCAGGCTCTAG

>gi|7682684|gb|AAF67343.1|AF240634_1 phosphoprotein associated with GEMs
[Homo sapiens] (SEQ ID NO:14)
15   MGPAGSLLGSGQMQITLWGSLAAVAIFFVITFLIFLCSSCDREKKPRQHSGDHENLMNVPSDKEMFSRSV
TSLATDAPASSEQNGALTNGDILSEDSTLTCMQHYEEVQTSASDLLDSQDSTGKPKCHQSRELPRIPPES
AVDTMLTARSVDGDQGLGMEGPYEVLKDSSSQENMVEDCLYETVKEIKEVAAAAHLEKGHSGKAKSTSAS
KELPGPQTEGKAEFAEYASVDRNKKCRQSVNVESILGNSCDPEEEAPPPVPVKLLDENENLQEKEGGEAE
ESATDTTSETNKRFSSLSYKSREEDPTLTEEEISAMYSSVNKPGQLVNKSGQSLTVPESTYTSIQGDPQR
20   SPSSCNDLYATVKDFEKTPNSTLPPAGRPSEEPEPDYEAIQTLNREEEKATLGTNGHHGLVPKENDYESI
SDLQQGRDITRL

SHP/PTP1C
25

>gi|35781|emb|X62055.1|HSPTP1C H. sapiens PTP1C mRNA for protein-tyrosine
phosphatase 1C (SEQ ID NO:15)
CAAGAAGACGGGGATTGAGGAGGCCTCAGGCGCCTTTGTCTACCTGCGGCAGCCGTACTATGCCACGAGG
GTGAATGCGGCTGACATTGAGAACCGAGTGTTGGAACTGAACAAGAAGCAGGAGTCCGAGGAGGAAGTGG
30   CTGATTACTGAGCGGTTCTTCCTCACCTGGCTTGGGCCACTGTGCACAGCTGTGCCGCTGGCTCAGCCCC
GCCCCCTGCGGCCCTCCGCCGTGGCTTCCCCCTCCCTACAGAGAGATGCTGTCCCGTGGGTGGTTTCACC
GAGACCTCAGTGGGCTGGATGCAGAGACCCTGCTCAAGGGCCGAGGTGTCCACGGTAGCTTCCTGGCTCG
GCCCAGTCGCAAGAACCAGGGTGACTTCTCGCTCTCCGTCAGGGTGGGGGATCAGGTGACCCATATTCGG
ATCCAGAACTCAGGGGATTTCTATGACCTGTATGGAGGGGAGAAGTTTGCGACTCTGACAGAGCTGGTGG
35   AGTACTACACTCAGCAGCAGGGTGTCCTGCAGGACCGCGACGGCACCATCATCCACCTCAAGTACCCGCT
GAACTGCTCCGATCCCACTAGTGAGAGGTGGTACCATGGCCACATGTCTGGCGGGCAGGCAGAGACGCTG
CTGCAGGCCAAGGGCGAGCCCTGGACGTTTCTTGTGCGTGAGAGCCTCAGCCAGCCTGGAGACTTCGTGC
TTTCTGTGCTCAGTGACCAGCCCAAGGCTGGCCCAGGCTCCCCGCTCAGGGTCACCCACATCAAGGTCAT
GTGCGAGGGTGGACGCTACACAGTGGGTGGTTTGGAGACCTTCGACAGCCTCACGGACCTGGTAGAGCAT
40   TTCAAGAAGACGGGGATTGAGGAGGCCTCAGGCGCCTTTGTCTACCTGCGGCAGCCGTACTATGCCACGA
GGGTGAATGCGGCTGACATTGAGAACCGAGTGTTGGAACTGAACAAGAAGCAGGAGTCCGAGGATACAGC
CAAGGCTGGCTTCTGGGAGGAGTTTGAGAGTTTGCAGAAGCAGGAGGTGAAGAACTTGCACCAGCGTCTG
GAAGGGCAGCGGCCAGAGAACAAGGGCAAGAACCGCTACAAGAACATTCTCCCCTTTGACCACAGCCGAG
TGATCCTGCAGGGACGGGACAGTAACATCCCCGGGTCCGACTACATCAATGCCAACTACATCAAGAACCA
45   GCTGCTAGGCCCTGATGAGAACGCTAAGACCTACATCGCCAGCCAGGGCTGTCTGGAGGCCACGGTCAAT
GACTTCTGGCAGATGGCGTGGCAGGAGAACAGCCGTGTCATCGTCATGACCACCCGAGAGGTGGAGAAAG
GCCGGAACAAATGCGTCCCATACTGGCCCGAGGTGGGCATGCAGCGTGCTTATGGGCCCTACTCTGTGAC
CAACTGCGGGGAGCATGACACAACCGAATACAAACTCCGTACCTTACAGGTCTCCCCGCTGGACAATGGA
GACCTGATTCGGGAGATCTGGCATTACCAGTACCTGAGCTGGCCCGACCATGGGGTCCCCAGTGAGCCTG
50   GGGGTGTCCTCAGCTTCCTGGACCAGATCAACCAGCGGCAGGAAAGTCTGCCTCACGCAGGGCCCATCAT
CGTGCACTGCAGCGCCGGCATCGGCCGCACAGGCACCATCATTGTCATCGACATGCTCATGGAGAACATC
TCCACCAAGGGCCTGGACTGTGACATTGACATCCAGAAGACCATCCAGATGGTGCGGGCGCAGCGCTCGG
GCATGGTGCAGACGGAGGCGCAGTACAAGTTCATCTACGTGGCCATCGCCCAGTTCATTGAAACCACTAA
GAAGAAGCTGGAGGTCCTGCAGTCGCAGAAGGGCCAGGAGTCGGAGTACGGGAACATCACCTATCCCCCA
55   GCCATGAAGAATGCCCATGCCAAGGCCTCCCGCACCTCGTCCAAACACAAGGAGGATGTGTATGAGAACC
TGCACACTAAGAACAAGAGGGAGGAGAAAGTGAAGAAGCAGCGGTCAGCAGACAAGGAGAAGAGCAAGGG
TTCCCTCAAGAGGAAGTGAGCGGTGCTGTCCTCAGGTGGCCATGCCTCAGCCCTGACCCTGTGGAAGCAT
TTCGCGATGGACAGACTCACAACCTGAACCTAGGAGTGCCCCATTCTTTTGTAATTTAAATGGCTGCATC
CCCCCCACCTCTCCCTGACCCTGTATATAGCCCAGCCAGGCCCCAGGCAGGGCCAACCCTTCTCCTCTTG
60   TAAATAAAGCCCTGGGATCACTGAAAAAAAAAAAAAA

>gi|35782|emb|CAA43982.1| protein-tyrosine phosphatase 1C [Homo sapiens]
(SEQ ID NO:16)
MLSRGWFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQGDFSLSVRVGDQVTHIRIQNSGDFYDLYGGEK
FATLTELVEYYTQQQGVLQDRDGTIIHLKYPLNCSDPTSERWYHGHMSGGQAETLLQAKGEPWTFLVRES
5   LSQPGDFVLSVLSDQPKAGPGSPLRVTHIKVMCEGGRYTVGGLETFDSLTDLVEHFKKTGIEEASGAFVY
LRQPYYATRVNAADIENRVLELNKKQESEDTAKAGFWEEFESLQKQEVKNLHQRLEGQRPENKGKNRYKN
ILPFDHSRVILQGRDSNIPGSDYINANYIKNQLLGPDENAKTYIASQGCLEATVNDFWQMAWQENSRVIV
MTTREVEKGRNKCVPYWPEVGMQRAYGPYSVTNCGEHDTTEYKLRTLQVSPLDNGDLIREIWHYQYLSWP
DHGVPSEPGGVLSFLDQINQRQESLPHAGPIIVHCSAGIGRTGTIIVIDMLMENISTKGLDCDIDIQKTI
10   QMVRAQRSGMVQTEAQYKFIYVAIAQFIETTKKKLEVLQSQKGQESEYGNITYPPAMKNAHAKASRTSSK
HKEDVYENLHTKNKREEKVKKQRSADKEKSKGSLKRK

CSK
15

>gi|4758077|ref|NM_004383.1| Homo sapiens c-src tyrosine kinase (CSK), mRNA
(SEQ ID NO:17)
TCCGGGGCGGCCCCCGGCAGCCAGCGCGACGTTCCAAAATCGAACCTCAGTGGCGGCGCTCGGAAGCGGA
ACTCTGCCGGGGCCGCGCCGGCTACATTGTTTCCTCCCCCGACTCCCTCCCGCCCCCTTCCCCCGCCTT
20   TCTTCCCTCCGCGACCCGGGCCGTGCGTCCGTCCCCCTGCCTCTGCCTGGCGGTCCCTCCTCCCCTCTCC
TTGCACCCATACCTCTTTGTACCGCACCCCCTGGGGACCCCTGCGCCCCTCCCCTCCCCCCTGACCGCAT
GGACCGTCCCGCAGGCCGCTGATGCCGCCCGCGGCGAGGTGGCCCGGACCGCAGTGCCCCAAGAGAGCTC
TAATGGTACCAAGTGACAGGTTGGCTTTACTGTGACTCGGGGACGCCAGAGCTCCTGAGAAGATGTCAGC
AATACAGGCCGCCTGGCCATCCGGTACAGAATGTATTGCCAAGTACAACTTCCACGGCACTGCCGAGCAG
25   GACCTGCCCTTCTGCAAAGGAGACGTGCTCACCATTGTGGCCGTCACCAAGGACCCCAACTGGTACAAAG
CCAAAAACAAGGTGGGCCGTGAGGGCATCATCCCAGCCAACTACGTCCAGAAGCGGGAGGGCGTGAAGGC
GGGTACCAAACTCAGCCTCATGCCTTGGTTCCACGGCAAGATCACACGGGAGCAGGCTGAGCGGCTTCTG
TACCCGCCGGAGACAGGCCTGTTCCTGGTGCGGGAGAGCACCAACTACCCCGGAGACTACACGCTGTGCG
TGAGCTGCGACGGCAAGGTGGAGCACTACCGCATCATGTACCATGCCAGCAAGCTCAGCATCGACGAGGA
30   GGTGTACTTTGAGAACCTCATGCAGCTGGTGGAGCACTACACCTCAGACGCAGATGGACTCTGTACGCGC
CTCATTAAACCAAAGGTCATGGAGGGCACAGTGGCGGCCCAGGATGAGTTCTACCGCAGCGGCTGGGCCC
TGAACATGAAGGAGCTGAAGCTGCTGCAGACCATCGGGAAGGGGGAGTTCGGAGACGTGATGCTGGGCGA
TTACCGAGGGAACAAAGTCGCCGTCAAGTGCATTAAGAACGACGCCACTGCCCAGGCCTTCCTGGCTGAA
GCCTCAGTCATGACGCAACTGCGGCATAGCAACCTGGTGCAGCTCCTGGGCGTGATCGTGGAGGAGAAGG
35   GCGGGCTCTACATCGTCACTGAGTACATGGCCAAGGGGAGCCTTGTGGACTACCTGCGGTCTAGGGGTCG
GTCAGTGCTGGGCGGAGACTGTCTCCTCAAGTTCTCGCTAGATGTCTGCGAGGCCATGGAATACCTGGAG
GGCAACAATTTCGTGCATCGAGACCTGGCTGCCCGCAATGTGCTGGTGTCTGAGGACAACGTGGCCAAGG
TCAGCGACTTTGGTCTCACCAAGGAGGCGTCCAGCACCCAGGACACGGGCAAGCTGCCAGTCAAGTGGAC
AGCCCCTGAGGCCCTGAGAGAGAAGAAATTCTCCACTAAGTCTGACGTGTGGAGTTTCGGAATCCTTCTC
40   TGGGAAATCTACTCCTTTGGGCGAGTGCCTTATCCAAGAATTCCCCTGAAGGACGTCGTCCCTCGGGTGG
AGAAGGGCTACAAGATGGATGCCCCCGACGGCTGCCCGCCCGCAGTCTATGAAGTCATGAAGAACTGCTG
GCACCTGGACGCCGCCATGCGGCCCTCCTTCCTACAGCTCCGAGAGCAGCTTGAGCACATCAAAACCCAC
GAGCTGCACCTGTGACGGCTGGCCTCCGCCTGGGTCATGGGCCTGTGGGGACTGAACCTGGAAGATCATG
GACCTGGTGCCCCTGCTCACTGGGCCCGAGCCTGAACTGAGCCCCAGCGGGCTGGCGGGCCTTTTTCCTG
45   CGTCCCAGCCTGCACCCCTCCGGCCCCGTCTCTCTTGGACCCACCTGTGGGGCCTGGGGAGCCCACTGAG
GGGCCAGGGAGGAAGGAGGCCACGGAGCGGGCGGCAGCGCCCCACCACGTCGGGCTTCCCTGGCCTCCCG
CCACTCGCCTTCTTAGAGTTTTATTCCTTTCCTTTTTTGAGATTTTTTTTCCGTGTGTTTATTTTTATT
ATTTTTCAAGATAAGGAGAAAGAAAGTACCCAGCAAATGGGCATTTTACAAGAAGTACGAATCTTATTTT
TCCTGTCCTGCCCGTGAGGTGGGGGGGACCGGGCCCCTCTCTAGGGACCCCTCGCCCCAGCCTCATTCCC
50   CATTCTGTGTCCCATGTCCCGTGTCTCCTCGGTCGCCCCGTGTTTGCGCTTGACCATGTTGCACTGTTTG
CATGCGCCCGAGGCAGACGTCTGTCAGGGGCTTGGATTTCGTGTGCCGCTGCCACCCGCCCACCCGCCTT
GTGAGATGGAATCGTAATAAACCACGCCATGAGGAAAAA

>gi|4758078|ref|NP_004374.1| c-src tyrosine kinase [Homo sapiens] (SEQ ID
55   NO:18)
MSAIQAAWPSGTECIAKYNFHGTAEQDLPFCKGDVLTIVAVTKDPNWYKAKNKVGREGIIPANYVQKREG
VKAGTKLSLMPWFHGKITREQAERLLYPPETGLFLVRESTNYPGDYTLCVSCDGKVEHYRIMYHASKLSI
DEEVYFENLMQLVEHYTSDADGLCTRLIKPKVMEGTVAAQDEFYRSGWALNMKELKLLQTIGKGEFGDVM
LGDYRGNKVAVKCIKNDATAQAFLAEASVMTQLRHSNLVQLLGVIVEEKGGLYIVTEYMAKGSLVDYLRS
60   RGRSVLGGDCLLKFSLDVCEAMEYLEGNNFVHRDLAARNVLVSEDNVAKVSDFGLTKEASSTQDTGKLPV
KWTAPEALREKKFSTKSDVWSFGILLWEIYSFGRVPYPRIPLKDVVPRVEKGYKMDAPDGCPPAVYEVMK

NCWHLDAAMRPSFLQLREQLEHIKTHELHL

nucleolin (NCL)

>gi|4885510|ref|NM_005381.1| Homo sapiens nucleolin (NCL), mRNA (SEQ ID
NO:19)
CTTCGGGTGTACGTGCTCCGGGATCTTCAGCACCCGCGGCCGCCATCGCCGTCGCTTGGCTTCTTCTGGA
CTCATCTGCGCCACTTGTCCGCTTCACACTCCGCCGCCATCATGGTGAAGCTCGCGAAGGCAGGTAAAAA
TCAAGGTGACCCCAAGAAAATGGCTCCTCCTCCAAAGGAGGTAGAAGAAGATAGTGAAGATGAGGAAATG
TCAGAAGATGAAGAAGATGATAGCAGTGGAGAAGAGGTCGTCATACCTCAGAAGAAAGGCAAGAAGGCTG
CTGCAACCTCAGCAAAGAAGGTGGTCGTTTCCCCAACAAAAAAGGTTGCAGTTGCCACACCAGCCAAGAA
AGCAGCTGTCACTCCAGGCAAAAAGGCAGCAGCAACACCTGCCAAGAAGACAGTTACACCAGCCAAAGCA
GTTACCACACCTGGCAAGAAGGGAGCCACACCAGGCAAAGCATTGGTAGCAACTCCTGGTAAGAAGGGTG
CTGCCATCCCAGCCAAGGGGGCAAAGAATGGCAAGAATGCCAAGAAGGAAGACAGTGATGAAGAGGAGGA
TGATGACAGTGAGGAGGATGAGGAGGATGACGAGGACGAGGATGAGGATGAAGATGAAATTGAACCAGCA
GCGATGAAAGCAGCAGCTGCTGCCCCTGCCTCAGAGGATGAGGACGATGAGGATGACGAAGATGATGAGG
ATGACGATGACGATGAGGAAGATGACTCTGAAGAAGAAGCTATGGAGACTACACCAGCCAAAGGAAAGAA
AGCTGCAAAAGTTGTTCCTGTGAAAGCCAAGAACGTGGCTGAGGATGAAGATGAAGAAGAGGATGATGAG
GACGAGGATGACGACGACGACGAAGATGATGAAGATGATGATGATGAAGATGATGAGGAGGAGGAAGAAG
AGGAGGAGGAAGAGCCTGTCAAAGAAGCACCTGGAAAACGAAAGAAGGAAATGGCCAAACAGAAAGCAGC
TCCTGAAGCCAAGAAACAGAAAGTGGAAGGCACAGAACCGACTACGGCTTTCAATCTCTTTGTTGGAAAC
CTAAACTTTAACAAATCTGCTCCTGAATTAAAAACTGGTATCAGCGATGTTTTTGCTAAAAATGATCTTG
CTGTTGTGGATGTCAGAATTGGTATGACTAGGAAATTTGGTTATGTGGATTTTGAATCTGCTGAAGACCT
GGAGAAAGCGTTGGAACTCACTGGTTTGAAAGTCTTTGGCAATGAAATTAAACTAGAGAAACCAAAAGGA
AAAGACAGTAAGAAAGAGCGAGATGCGAGAACACTTTTGGCTAAAAATCTCCCTTACAAAGTCACTCAGG
ATGAATTGAAAGAAGTGTTTGAAGATGCTGCGGAGATCAGATTAGTCAGCAAGGATGGGAAAAGTAAAGG
GATTGCTTATATTGAATTTAAGACAGAAGCTGATGCAGAGAAAACCTTTGAAGAAAAGCAGGGAACAGAG
ATCGATGGGCGATCTATTTCCCTGTACTATACTGGAGAGAAAGGTCAAAATCAAGACTATAGAGGTGGAA
AGAATAGCACTTGGAGTGGTGAATCAAAAACTCTGGTTTTAAGCAACCTCTCCTACAGTGCAACAGAAGA
AACTCTTCAGGAAGTATTTGAGAAAGCAACTTTTATCAAAGTACCCAGAACCAAAATGGCAAATCTAAA
GGGTATGCATTTATAGAGTTTGCTTCATTCGAAGACGCTAAAGAAGCTTTAAATTCCTGTAATAAAAGGG
AAATTGAGGGCAGAGCAATCAGGCTGGAGTTGCAAGGACCCAGGGGATCACCTAATGCCAGAAGCCAGCC
ATCCAAAACTCTGTTTGTCAAAGGCCTGTCTGAGGATACCACTGAAGAGACATTAAAGGAGTCATTTGAC
GGCTCCGTTCGGGCAAGGATAGTTACTGACCGGGAAACTGGGTCCTCCAAAGGGTTTGGTTTTGTAGACT
TCAACAGTGAGGAGGATGCCAAGGAGGCCATGGAAGACGGTGAAATTGATGGAAATAAAGTTACCTTGGA
CTGGGCCAAACCTAAGGGTGAAGGTGGCTTCGGGGGTCGTGGTGGAGGCAGAGGCGGCTTTGGAGGACGA
GGTGGTGGTAGAGGAGGCCGAGGAGGATTTGGTGGCAGAGGCCGGGGAGGCTTTGGAGGGCGAGGAGGCT
TCCGAGGAGGCAGAGGAGGAGGAGGTGACCACAAGCCACAAGGAAAGAAGACGAAGTTTGAATAGCTTCT
GTCCCTCTGCTTTCCCTTTTCCATTTGAAAGAAAGGACTCTGGGGTTTTTACTGTTACCTGATCAATGAC
AGAGCCTTCTGAGGACATTCCAAGACAGTATACAGTCCTGTGGTCTCCTTGGAAATCCGTCTAGTTAACA
TTTCAAGGGCAATACCGTGTTGGTTTTGACTGGATATTCATATAAACTTTTTAAAGAGTTGAGTGATAGA
GCTAACCCTTATCTGTAAGTTTTGAATTTATATTGTTTCATCCCATGTACAAAACCATTTTTTCCTAC

>gi|4885511|ref|NP_005372.1| nucleolin [Homo sapiens] (SEQ ID NO:20)
MVKLAKAGKNQGDPKKMAPPPKEVEEDSEDEEMSEDEEDDSSGEEVVIPQKKGKKAAATSAKKVVVSPTK
KVAVATPAKKAAVTPGKKAAATPAKKTVTPAKAVTTPGKKGATPGKALVATPGKKGAAIPAKGAKNGKNA
KKEDSDEEEDDDSEEDEEDDEDEDEDEDEIEPAAMKAAAAAPASEDEDDEDDEDDEDDDDDEEDDSEEEA
METTPAKGKKAAKVVPVKAKNVAEDEDEEEDDEDEDDDDDEDDEDDDDEDDEEEEEEEEEEPVKEAPGKR
KKEMAKQKAAPEAKKQKVEGTEPTTAFNLFVGNLNFNKSAPELKTGISDVFAKNDLAVVDVRIGMTRKFG
YVDFESAEDLEKALELTGLKVFGNEIKLEKPKGKDSKKERDARTLLAKNLPYKVTQDELKEVFEDAAEIR
LVSKDGKSKGIAYIEFKTEADAEKTFEEKQGTEIDGRSISLYYTGEKGQNQDYRGGKNSTWSGESKTLVL
SNLSYSATEETLQEVFEKATFIKVPQNQNGKSKGYAFIEFASFEDAKEALNSCNKREIEGRAIRLELQGP
RGSPNARSQPSKTLFVKGLSEDTTEETLKESFDGSVRARIVTDRETGSSKGFGFVDFNSEEDAKEAMEDG
EIDGNKVTLDWAKPKGEGGFGGRGGGRGGFGGRGGGRGGRGGFGGRGRGGFGGRGGFRGGRGGGGDHKPQ
GKKTKFE

SLAP

>gi|5804883|emb|AJ238592.1|HSA238592 Homo sapiens SLAP gene promoter region
(SEQ ID NO:21)
```
AATTCTCTCCCCGTCTGTCCTCTGAATATAGGTTATACATACACGAAATTTACAATAGATGTTGGCTAAA
GTGAGTAATTCAGCAAATACGTAATGAACATTGTTTTGTTCTGGGCCTGGAAAACAGAGAAAAAGAGTG
TTTAGCTTTGAACTCCTTGAGGGTAGGGCCTAGTCTTGCCTGGCACATAGTACAGTTACATACATAGCGT
CTGGCACACAGTGAGTGCTAACAAAATTTGTTGAATGAATAACTGAATGAATGTATGGAGCACAAAATGC
ATGCTTGAATAAATGACAATATTCTTGGACTCGAGACTTCTTGATGGATTAGAAGAAAAAGGTGTGTGAA
TCACTGAGGAACAGTATGGAAAATGGTTGAGAGAGGAATGGTTAGGGTGGTGGAGGAGCAGGGAAGACAT
GGAAGCGAACTCACAATACAGTCCTCCCTGGCTGGGCCTCTGTTGTGAGAAGTCCACCTCCTGCTCTTCC
TCTACTGTCCTTAGGCAGCTCGTGGCTCACATGCTCCCACCCCAGTTCTTGACAGCAAACTGCTCCTCCT
TTAAGATGCAGCTCACAGGCTGCCCATACTGTAGCTTCCAAGCCCTAATTCCCAGGGACTCCTGAACAGA
CCACAGACCAGTTCCTCTTCTTGGGTCACGTGTATGATGGTGGAGAGGAAGGATGGGTGACACAGCCACC
CAAAGGAGCCTTGCTACACAGCAGCTTCTTAGATGTGTACCAGAGAGCACAGGCAGACGTGCTGCAGAGC
TGAGTCCACACAAAAGCAGGAAAACCAACTGCCCTCAGGGCAGCACTTTACAACCCAAGAAAGAACAACC
TTCACACGTGTCTTATTTTGGGATGGGTATGATCCAGTGGTTCTCAACTGGGGGCAATTTTTGCTCCCTA
GGGCACATTTGACAATTTCTGGGACATTTTTGGTTGTCATGACTTGGAGGGGTGCGGGGGGCTCCTGCTG
TCTAGGGGGTCAGGTCCAGAGATGCTGTGGCACATCCTCCAATGCACAGGTCAGCCCTGCAACTCTGCCG
CACACACACACATACAAACACACACACACACACACACACACATACACACCAGAGAATGCCCTGATCCA
AAATGTCCAAAGTGCTGAGGCTGAGACCCCGCTTGGGGTTACTGAAAACAGAGGCCTAAGATTTGAGTCA
CCTGTCTAGCGTGGCATCGGGCTTCAGCCCTGCCGTCTCTGGTTGGGCCCAGGGCCAACACCTCCATTCC
ATCCCCCTGCTCTGGCACAGCCCAGGCTTTGGGATAGCACAGTGCCCAGTATGAGGAACACTTTCTCAAA
CCTACTAAGTTTCCTCGGTTACTCCAGGAGGCGAAAGGGAGGGTCATAAACACCCTGGGCTGGGGGTCGG
GGGAGGTTAGCAGCAACCCTCCAGTCACCCAATGAGCCACTGGACAAAAACAGAGAAGAGCCTGCGGCCC
TGACCTCCAGGGCTCTGCATCCTGCCTTTCTCCTTCTTCTGTAGCTTCTCCTGTGAACGGGACCCCTTGA
ACAGGAACCCACTCTATGGGCCTCTCTCTTGTCTGCCTCTGCCTCTGACTTGAACATCTTTAGGGCTGAT
TGTCTTTCCCAAAAACTTTTAAAGCTTGTGTGAAACCAGCAGTCCCTTCTGCGCACACCTCTCATGCCCT
GGCTTGTCCCAGGGCTGGCTTCCTTTCCTCCTACAGTTCTCAGCCTAATGATATCTCTAGAGAGAGCCCA
TCCACCCGACGACTGGCCACCTGTGCAAAGTAGCACATCTGCGCCCCTCGGTCCCATCAACCCATTGAGT
TTCTTTGATGGGTCGCCATAATTGATAGTCACATATTCAGTCTTTCCCATAGGACTCCTCCTAGGGGAAG
TAGTAGAATCTGCTGTGTAGTACATATTCAAATAACCGCCACCCATAATGTTCTTTCTAGTTCTTTTCC
CGTGACCCACACTTTGCACCAAGTCGCCATGTCTCTGTAGTCCCCTTTAAATTAAGTTAAGTACAGGGAA
GGGCAGGGCTTTTGCAGTTGTTTTTTTTTTTTTTTTTTTTTACAGCTAAATCCCAGTTTGTGGAATATGC
CTTGCTCAATAAAGTCTTACTAAAGAAATGACAGGGAATAAATTCAGAGCTTCATCTGTACTAGATTCTT
TCTTGGTAAGTGTCTATTAAAAAGTCATCACAAAACTCATACCTCTGACTATCAGTCTAAAGCTCTTTTA
CTCCTCTGTGTGAATTTGGGGGAAAATCTCCACTTCTGCATATTTCTGCCTATCCCGTCTTCCCTGTAAA
GCAAACTGCCCACTCAGTCTTTTGCAATAGTTCCTTTGTACCTTCCCAGCATGCTCTTCTTGAGCGGGTG
TTTTCACAAGTTGTCTAAGGAACCATCTCTCTAGTGCAGAGTGCCTGCCTGCTAAATTTAGAAGCAGCCC
TGAGGGTTTTGATTCATCTTTGCAAACTACCCCAAGTGAGGTTTCCAAATGACTGTCCCAGGGTGGCTGT
GTCAGCCTCACTGGACTGGGGCCTGGGTCACCGGCCGTAGATTCTGGCCCTGGCTCTTCTTTGGGACTAG
CTGTGTGACTGTGGACAGATCACTCAAGTCTCTGATGCTCTGTCCACTCTGTGAGGTGTGAGGGGCTGTG
AGGGGCTCTGAGGGCCATCCCAGTGTGGTTACCTGGCATTCCAAGCTCAAGCAGGCAATTCCCAGGCAAT
CTTGGGGTCAGCATTCCAGAAATTTTACCTGTTTCATCAATAAGCTCTGAAGCCCCATGTGCCTGAGTTT
CTCTGCTTGAGACTCACGTACTGTAAGTAGGGAGATTTTATATGCAATATTTACATAATGTATAACATAT
ACAACATACATTTTATAATATATAACATAAAGACACGTATTTTAACATCAAACTCACCAAAAACTTGC
AAGAGTAGAACAATGCACACCTATATCCCCTTCACCTAGGCTCACCAAGTATTAACATTTTGTCCCGTTT
GCCTTCTGTCTCCCTCATTCTCTTCCTGTGTCTTGATATATAGATGTAGATGTATTAGGTTGGTGCCACA
ATGATTGTGGTTTTTGCCATTACTTCTATGGCAGAAACCACAATAACTTTTGTGCCAATATGATGGCTGT
AGATATAGGGATGCAGCTTTTCTTTTAACAAACTCTTTGAAAGATATAGCCGTCATGATATGAGACATCC
TCTTGGTTACCGCATGTCCCCTAAGAACTAGACGTTCTCCCGCATAACCACACTATTATTACCATGTCCA
AGAAACTTAAGACTGGAACAGTATTATCAACTGTCAGTATATATTCAAACAACTGCCATCTCATAATGT
TCTTTCTAGTTTTTTTCCCTTGACGCACACATTGGCACCAAGTCATTGTGTCTCTTTAGTCTCCTTTAAT
CTCAAAGCATTGCTCTGGGGTCCATTCGGCAGTAACAGAAGGAAGCTGCAAAAATTGGGAAACCGAAATC
AGAGAAGCTGCAGCTCTTAGCCACAGCCTTCGTGGGCCCAGAGAGCTCTGGGGATGTACAAGAGCAAATT
CACAGCCCCATTGATTTGCCTTTTAATTAGAGAAAGAAGACCAAGCTTCTGAAATCCCAGAATCTAGAGT
TAAGCTGGCCATATTCTCAAATGCTTTGTAAAAAAATTAATAACCCATCTATCTTCCCAGTCCTTGCAGA
GAGAAGCTTGCTCAGAGTGGTTACTGAAAACAGGAACCTGCCCTTCTGCAAGCCCTGTGCCTTCCTCGTT
CCTAGCAGAGGAAGTATGCGGCCAGTTTAGCAACTGCAGTTCTTCTGTGTAGCATGGGCTGCTTCTCCAC
AACCGCGAGAGGCCCGGCTGCCTGTTTATTCTGGAGCAGCAGCAGTTGACAGACGGAGAGGCAGTGACTC
TGCACCCAGCAATCCCTCTGCAGACAGCTCCCACAGGAGTCATGCCCAGTGTGTGAAGGTATTTTGGGGG
GATGTGGCAGAGATGGGGTGGGGAGAGACTTCAGGGCCTCTTCAACACGCAGTTGGCAGGGGTTGACTTT
CTGGAGTCAGAAAGGAGCTCCCTGAGAACAGGTGGGGCAATGTCTTGAGAGAAGCCCAGCCCCCAACCTG
```

GTGCCCAGCACACTGCAGTGGCTCAGTGGGTGTTGGATGCATAATTCTGTGGAGAGAAGGGCTGGAGGAG
CAGGGGCAGGATGCTGCCAGAGGAGACATTCTCGCCAAAAAGGCTTGTACTTAGCAGGCTTCTTACAAAG
GACAATTGGGATTTGCCATCAGAATT

5

>gi|5803170|ref|NM_006748.1| Homo sapiens Src-like-adaptor (SLA), mRNA (SEQ
ID NO:22)
GGCTCTGGGCATCACCAGCGGCCCCAGGGAAAAAGAAAGAAATGGGAAACAGCATGAAATCCACCCCTGC
GCCTGCCGAGAGGCCCCTGCCCAACCCGGAGGGACTGGATAGCGACTTCCTTGCCGTGCTAAGTGACTAC
10 CCGTCTCCTGACATCAGCCCCCCGATATTCCGCCGAGGGGAGAAACTGCGTGTGATTTCTGATGAAGGGG
GCTGGTGGAAAGCTATTTCTCTTAGCACTGGTCGAGAGAGTTACATCCCTGGAATATGTGTGGCCAGAGT
TTACCATGGCTGGCTGTTTGAGGGCCTGGGCAGAGACAAGGCCGAGGAGCTGCTGCAGCTGCCAGACACA
AAGGTCGGCTCCTTCATGATCAGAGAGAGTGAGACCAAGAAAGGGTTTTACTCACTGTCGGTGAGACACA
GGCAGGTAAAGCATTACCGCATTTTCCGTCTGCCGAACAACTGGTACTACATTTCCCCGAGGCTCACCTT
15 CCAGTGCCTGGAGGACCTGGTGAACCACTATTCTGAGGTGGCTGATGGCCTGTGCTGTGTGCTCACCACG
CCCTGCCTGACACAAAGCACGGCTGCCCCAGCAGTGAGGGCCTCCAGCTCACCTGTCACCTTGCGTCAGA
AGACTGTGGACTGGAGGAGAGTGTCCAGACTGCAGGAGGACCCCGAGGGAACAGAGAACCCGCTTGGGGT
AGACGAGTCCCTTTTCAGCTATGGCCTTCGAGAGAGCATTGCCTCTTACCTGTCCCTGACCAGTGAGGAC
AACACCTCCTTTGATCGAAAGAAGAAAAGCATCTCCCTGATGTATGGTGGCAGCAAGAGAAAGAGCTCAT
20 TCTTCTCATCACCACCTTACTTTGAGGACTAGCCAAGAACAGACACAATGGTTCATGCCCAAAAGGAACA
GAAGTTCCAACTATTGCCTGGGATCTTGCGAAAAGCGAGGTTCCCTGATCCCTGGGAGCCTCACGTATTT
TAGAAGCCAAGAGAAGCCACATGGAGACTCAAATTCGCATCTTCTCTATCCACATCATGACCAAAGGAAC
CCCTCCCTGGTGTCTGATCAGGGCTGTGGCATCACAAAACATTGGATCATGACATGTCGGGCGATGCTTG
GAAAAGCCCAGCATGTATGTATGCACACATTGTGTGTGTGGGAAGGACAAAGCCACTCTCACAAGAAAGG
25 GCACCAGGACTGCTCTCCAAGGAACTGGACCTGTCCAGACAGTTACACTCCAAGGTCATTGGAGAGAACT
TCTGTATGGGCAAGCCTGAGAGGGAGAGGAAACAAAAGCTGTGTCCTGGCAGAAGGTCTGGGTTTGCAGA
TGGGTGCCCTGAATGGAACTACTTTAACTAATCCATAGGGACTTCTGGTATGCTTTCCTCTCTTTTTAAA
GGAACTTCGTGACACTAAACATTAGCCCAAAGGACTTCTTAGCCTTCAATTGGGAGATACCTTTGGTCTG
CTCCTGCACCAAAGCCATATGGGTGGAAGTCAGTTGGCCTCCCTGGTTCTGCAGAGGGCCAGAAGAATGA
30 GAGAGAGGAAGACTGCTGGCAGGGAAATCGAGGAGGCGAGACTAGAACTGCACCAGCTTCCCTGATGTCT
GCAGCCATGGCTTTGCAGCGCAAACAGAACTTCTCTGGGATGCTGGGATTCTTGCCTGTATGAATGCATC
AAGTATTCATTTATTGCCCGAATAGGCATTGCATTAAGTCCTCTGTTAGGTGTCAGGCAAGCCAAAAAAA
AAAAAAAGATGCTAAGTCCTAACCCCCAACAGAAGTGTTCACAGTGTAGACGGGAAAAAATGTATAAACA
AATGTGTAAAAAGAGAAATCAGCTCATGGCTTAGGATGGAATTAGAGACAGGTGAGGGACACTCAGGAGC
35 TCATTTTCCAGCTGCTCTTCAGAGTGGAAGGGCTGGCTGGATCGGGTAGGTAAGAATAGCTGGATTTTTT
AGAAAGAAATGGATACAGTCTAAAGAATTAACTCACCCGGTACTTTATTCTAAGAAGGGTCTGGCATCC
ATATGAGGAAAAATGCTCAGCTCCAGGAAAGATGGGGAGTCCAAGTGGATTAATGATGTCATGCATAATT
TTAAGAGACAAGGGAGAAAACACAATGTATAGCCAGAGAAGGAGAAGCTCCCATCCAAATCCTACTAGGA
AGAGAGTGGGCTGCAGATGAATCTGTGACTCATGTTTCCCTGTTTCAAAGGGATCCTGGGGAAGGAGGGG
40 AACATGCTTGCAGTATCTCTCCCTGTCTGTCTGCTCACATAAGCATTCCGTCCATCTAAGCTCATCGTGC
TACTGGTATGTGTATGTGCAGTTACACAGTTTCCTGTATCATAAATCCTAGTGTGTTTATACAAGGAGAC
ATCTGTGGTTTCCCCAACCGTTCCAAAAGGCTATTTCAAAGGAACCAGCCCACGTATGAGAAATGAATGT
AACACTGTGGACATTGACTTCCCGCATAAGGCAGGGTGACCCCCTGAACTCCAGATGTTTGCACAGTATC
TTATGTGTTGTTTTCCGTTGTGACGAATGTGATTGGAACATTTGGGGAGCACCCAGAGGGATTTTTCAGT
45 GGGAAGCATTACACTTTGCTAAATCATGTATTTATTCCTGATTAAAACAAACCTAATAAATATTTAACCC
TTGGC

>gi|5803171|ref|NP_006739.1| Src-like-adaptor; Src-like-adapter [Homo
sapiens] (SEQ ID NO:23)
50 MGNSMKSTPAPAERPLPNPEGLDSDFLAVLSDYPSPDISPPIFRRGEKLRVISDEGGWWKAISLSTGRES
YIPGICVARVYHGWLFEGLGRDKAEELLQLPDTKVGSFMIRESETKKGFYSLSVRHRQVKHYRIFRLPNN
WYYISPRLTFQCLEDLVNHYSEVADGLCCVLTTPCLTQSTAAPAVRASSSPVTLRQKTVDWRRVSRLQED
PEGTENPLGVDESLFSYGLRESIASYLSLTSEDNTSFDRKKKSISLMYGGSKRKSSFFSSPPYFED

55

PAK2

>gi|4505598|ref|NM_002577.1| Homo sapiens p21 (CDKN1A)-activated kinase 2
(PAK2), mRNA (SEQ ID NO:24)
60 GACCTTGGCTTGCCCGGGGCCATTTCATAATTCTGAATCATGTCTGATAACGGAGAACTGGAAGATAAGC
CTCCAGCACCTCCTGTGCGAATGAGCAGCACCATCTTTAGCACTGGAGGCAAAGACCCTTTGTCAGCCAA

TCACAGTTTGAAACCTTTGCCCTCTGTTCCAGAAGAGAAAAAGCCCAGGCATAAAATCATCTCCATATTC
TCAGGCACAGAGAAAGGAGTAAAAAGAAAGAAAAGGAACGGCCAGAAATTTCTCCTCCATCTGATTTTG
AGCACACCATCCATGTTGGCTTTGATGCTGTTACTGGAGAATTCACTGGCATGCCAGAACAGTGGGCTCG
ATTACTACAGACCTCCAATATCACCAAACTAGAGCAAAAGAAGAATCCTCAGGCTGTGCTGGATGTCCTA
AAGTTCTACGACTCCAACACAGTGAAGCAGAAATATCTGAGCTTTACTCCTCCTGAGAAAGATGGCCTTC
CTTCTGGAACGCCAGCACTGAATGCCAAGGGAACAGAAGCACCCGCAGTAGTGACAGAGGAGGAGGATGA
TGATGAAGAGACTGCTCCTCCCGTTATTGCCCCGCGACCGGATCATACGAAATCAATTTACACACGGTCT
GTAATTGACCCTGTTCCTGCACCAGTTGGTGATTCACATGTTGATGGTGCTGCCAAGTCTTTAGACAAAC
AGAAAAAGAAGCCTAAGATGACAGATGAAGAGATTATGGAGAAATTAAGAACTATCGTGAGCATAGGTGA
CCCTAAGAAAAAATATACAAGATATGAAAAAATTGGACAAGGGGCTTCTGGTACAGTTTTCACTGCTACT
GACGTTGCACTGGGACAGGAGGTTGCTATCAAACAAATTAATTTACAGAAACAGCCAAAGAAGGAACTGA
TCATTAACGAGATTCTGGTGATGAAAGAATTGAAAAATCCCAACATCGTTAACTTTTTGGACAGTTACCT
GGTAGGAGATGAATTGTTTGTGGTCATGGAATACCTTGCTGGGGGGTCACTCACTGATGTGGTAACAGAA
ACAGCTTGCATGGATGAAGCACAGATTGCTGCTGTATGCAGAGAGTGTTTACAGGCATTGGAGTTTTTAC
ATGCTAATCAAGTGATCCACAGAGACATCAAAAGTGACAATGTACTTTTGGGAATGGAAGGATCTGTTAA
GCTCACTGACTTTGGTTTCTGTGCCCAGATCACCCCTGAGCAGAGCAAACGCAGTACCATGGTCGGAACG
CCATACTGGATGGCACCAGAGGTGGTTACACGGAAAGCTTATGGCCCTAAAGTCGACATATGGTCTCTGG
GTATCATGGCTATTGAGATGGTAGAAGGAGAGCCTCCATACCTCAATGAAAATCCCTTGAGGGCCTTGTA
CCTAATAGCAACTAATGGAACCCCAGAACTTCAGAATCCAGAGAAACTTTCCCCAATATTTCGGGATTTC
TTAAATCGATGTTTGGAAATGGATGTGGAAAAAAGGGGTTCAGCCAAAGAATTATTACAGCATCCTTTCC
TGAAACTGGCCAAACCGTTATCTAGCTTGACACCACTGATCATGGCAGCTAAAGAAGCAATGAAGAGTAA
CCGTTAACATCACTGCTGTGGGCTCATACTCTTTTTTCCATTTTCTACAAGAAGCCTTTTAGTATATGAA
AATGATGACTCTGTTGGGGGTTTAAAGAAATGGTCTGCATAACCTGAATGAAAGAAGGAAATGACTATTC
TCTGAAGACAACCAAGAGAAAATTGGAAAAGACAAGGTATGACTTTGTTATGAACCCCTGCTTTTAGGGG
TCCAGGAAGGGATTTGTGGGACTTGAATTCACTAGGCTTAGGTCTTTCAGGAAACAGGCTATCAGGGGCA
TTTATCATGTGTGAGATTGGATTCTACTTGGGTGATTTGGTGGATAGACCCATGAATGGCCCCTGGGGGGT
TTTCAATCTTGGATTGGAGGTGGGGGTTTCAGAGTGTTGCCACGTCTAGCTCCTCTCCC

>gi|4505599|ref|NP_002568.1| p21 (CDKN1A)-activated kinase 2; novel serine
kinase; hPAK65 [Homo sapiens] (SEQ ID NO:25)
MSDNGELEDKPPAPPVRMSSTIFSTGGKDPLSANHSLKPLPSVPEEKKPRHKIISIFSGTEKGSKKKEKE
RPEISPPSDFEHTIHVGFDAVTGEFTGMPEQWARLLQTSNITKLEQKKNPQAVLDVLKFYDSNTVKQKYL
SFTPPEKDGLPSGTPALNAKGTEAPAVVTEEEDDDEETAPPVIAPRPDHTKSIYTRSVIDPVPAPVGDSH
VDGAAKSLDKQKKKPKMTDEEIMEKLRTIVSIGDPKKKYTRYEKIGQGASGTVFTATDVALGQEVAIKQI
NLQKQPKKELIINEILVMKELKNPNIVNFLDSYLVGDELFVVMEYLAGGSLTDVVTETACMDEAQIAAVC
RECLQALEFLHANQVIHRDIKSDNVLLGMEGSVKLTDFGFCAQITPEQSKRSTMVGTPYWMAPEVVTRKA
YGPKVDIWSLGIMAIEMVEGEPPYLNENPLRALYLIATNGTPELQNPEKLSPIFRDFLNRCLEMDVEKRG
SAKELLQHPFLKLAKPLSSLTPLIMAAKEAMKSNR

TCPTP/PTPN2

>gi|18104978|ref|NM_002828.2| Homo sapiens protein tyrosine phosphatase,
non-receptor type 2 (PTPN2), transcript variant 1, mRNA (updated 1/10/02)
(SEQ ID NO:26)
GCTCGGGCGCCGAGTCTGCGCGCTGACGTCCGACGCTCCAGGTACTTTCCCCACGGCCGACAGGGCTTGG
CGTGGGGGCGGGGCGCGGCGCGCAGCGCGCATGCGCCGCAGCGCCAGCGCTCTCCCCGGATCGTGCGGGG
CCTGAGCCTCTCCGCCGGCGCAGGCTCTGCTCGCGCCAGCTCGCTCCCGCAGCCATGCCCACCACCATCG
AGCGGGAGTTCGAAGAGTTGGATACTCAGCGTCGCTGGCAGCCGCTGTACTGGAAATTCGAAATGAGTC
CCATGACTATCCTCATAGAGTGGCCAAGTTTCCAGAAAACAGAAATCGAAACAGATACAGAGATGTAAGC
CCATATGATCACAGTCGTGTTAAACTGCAAAATGCTGAGAATGATTATATTAATGCCAGTTTAGTTGACA
TAGAAGAGGCACAAAGGAGTTACATCTTAACACAGGGTCCACTTCCTAACACATGCTGCCATTTCTGGCT
TATGGTTTGGCAGCAGAAGACCAAAGCAGTTGTCATGCTGAACCGCATTGTGGAGAAGAATCGGTTAAA
TGTGCACAGTACTGGCCAACAGATGACCAAGAGATGCTGTTTAAAGAAACAGGATTCAGTGTGAAGCTCT
TGTCAGAAGATGTGAAGTCGTATTATACAGTACATCTACTACAATTAGAAAATATCAATAGTGGTGAAAC
CAGAACAATATCTCACTTTCATTATACTACCTGGCCAGATTTTGGAGTCCCTGAATCACCAGCTTCATTT
CTCAATTTCTTGTTTAAAGTGAGAGAATCTGGCTCCTTGAACCCTGACCATGGGCCTGCGGTGATCCACT
GTAGTGCAGGCATTGGGCGCTCTGGCACCTTCTCTCTGGTAGACACTTGTCTTGTTTTGATGGAAAAAGG
AGATGATATTAACATAAAACAAGTGTTACTGAACATGAGAAAATACCGAATGGGTCTTATTCAGACCCCA
GATCAACTGAGATTCTCATACATGGCTATAATAGAAGGAGCAAAATGTATAAAGGGAGATTCTAGTATAC
AGAAACGATGGAAAGAACTTTCTAAGGAAGACTTATCTCCTGCCTTTGATCATTCACCAAACAAAATAAT

```
GACTGAAAAATACAATGGGAACAGAATAGGTCTAGAAGAAGAAAAACTGACAGGTGACCGATGTACAGGA
CTTTCCTCTAAAATGCAAGATACAATGGAGGAGAACAGTGAGAGTGCTCTACGGAAACGTATTCGAGAGG
ACAGAAAGGCCACCACAGCTCAGAAGGTGCAGCAGATGAAACAGAGGCTAAATGAGAATGAACGAAAAAG
AAAAAGGTGGTTATATTGGCAACCTATTCTCACTAAGATGGGGTTTATGTCAGTCATTTTGGTTGGCGCT
```
5
```
TTTGTTGGCTGGAGACTGTTTTTTTCAGCAAAATGCCCTATAAACAATTAATTTTGCCCAGCAAGCTTCTG
CACTAGTAACTGACAGTGCTACATTAATCATAGGGGTTTGTCTGCAGCAAACGCCTCATATCCCAAAAAC
GGTGCAGTAGAATAGACATCAACCAGATAAGTGATATTTACAGTCACAAGCCCAACATCTCAGGACTCTT
GACTGCAGGTTCCTCTGAACCCCAAACTGTAAATGGCTGTCTAAAATAAAGACATTCATGTTTGTTAAAA
ACTGGTAAATTTTGCAACTGTATTCATACATGTCAAACACAGTATTTCACCTGACCAACATTGAGATATC
```
10
```
CTTTATCACAGGATTTGTTTTTGGAGGCTATCTGGATTTTAACCTGCACTTGATATAAGCAATAAATATT
GTGGTTTTATCTACGTTATTGGAAAGAAATGACATTTAAATAATGTGTGTAATGTATAATGTACTATTG
ACATGGGCATCAACACTTTTATTCTTAAGCATTTCAGGGTAAATATATTTTATAAGTATCTATTTAATCT
TTTGTAGTTAACTGTACTTTTTAAGAGCTCAATTTGAAAAATCTGTTACTAAAAAAAAAAATTGTATGTC
GATTGAATTGTACTGGATACATTTTCCATTTTTCTAAAAAGAAGTTTGATATGAGCAGTTAGAAGTTGGA
```
15
```
ATAAGCAATTTCTACTATATATTGCATTTCTTTTATGTTTTACAGTTTTCCCCATTTTAAAAAGAAAAGC
AAACAAAGAAACAAAAGTTTTTCCTAAAAATATCTTTGAAGGAAATTCTCCTTACTGGGATAGTCAGGT
AAACAGTTGGTCAAGACTTTGTAAAGAAATTGGTTTCTGTAAATCCCATTATTGATATGTTTATTTTTCA
TGAAAATTTCAATGTAGTTGGGGTAGATTATGATTTAGGAAGCAAAGTAAGAAGCAGCATTTTATGATT
CATAATTTCAGTTTACTAGACTGAAGTTTTGAAGTAAACACTTTTCAGTTTCTTTCTACTTCAATAAATA
```
20
```
GTATGATTATATGCAAACCTTAAAAAA
```

>gi|4506290|ref|NM_002828.1| Homo sapiens protein tyrosine phosphatase, non-receptor type 2 (PTPN2), mRNA (SEQ ID NO:27)

```
GGGGGGCCTGAGCCTCTCCGCCGGCGCAGGCTCTGCTCGCGCCAGCTCGCTCCCGCAGCCATGCCCACCA
```
25
```
CCATCGAGCGGGAGTTCGAAGAGTTGGATACTCAGCGTCGCTGGCAGCCGCTGTACTTGGAAATTCGAAA
TGAGTCCCATGACTATCCTCATAGAGTGGCCAAGTTTCCAGAAAACAGAAATCGAAACAGATACAGAGAT
GTAAGCCCATATGATCACAGTCGTGTTAAACTGCAAAATGCTGAGAATGATTATATTAATGCCAGTTTAG
TTGACATAGAAGAGGCACAAAGGAGTTACATCTTAACACAGGGTCCACTTCCTAACACATGCTGCCATTT
CTGGCTTATGGTTTGGCAGCAGAAGACCAAAGCAGTTGTCATGCTGAACCGCATTGTGGAGAAGAATCG
```
30
```
GTTAAATGTGCACAGTACTGGCCAACAGATGACCAAGAGATGCTGTTTAAAGAAACAGGATTCAGTGTGA
AGCTCTTGTCAGAAGATGTGAAGTCGTATTATACAGTACATCTACTACAATTAGAAAATATCAATAGTGG
TGAAACCAGAACAATATCTCACTTTCATTATACTACCTGGCCAGATTTTGGAGTCCCTGAATCACCAGCT
TCATTTCTCAATTTCTTGTTTAAAGTGAGAGAATCTGGCTCCTTGAACCCTGACCATGGGCCTGCGGTGA
TCCACTGTAGTGCAGGCATTGGGCGCTCTGGCACCTTCTCTCTGGTAGACACTTGTCTTGTTTTGATGGA
```
35
```
AAAAGGAGATGATATTAACATAAAACAAGTGTTACTGAACATGAGAAAATACCGAATGGGTCTTATTCAG
ACCCCAGATCAACTGAGATTCTCATACATGGCTATAATAGAAGGAGCAAAATGTATAAAGGGAGATTCTA
GTATACAGAAACGATGGAAAGAACTTTCTAAGGAAGACTTATCTCCTGCCTTTGATCATTCACCAAACAA
AATAATGACTGAAAAATACAATGGGAACAGAATAGGTCTAGAAGAAGAAAAACTGACAGGTGACCGATGT
ACAGGACTTTCCTCTAAAATGCAAGATACAATGGAGGAGAACAGTGAGAGTGCTCTACGGAAACGTATTC
```
40
```
GAGAGGACAGAAAGGCCACCACAGCTCAGAAGGTGCAGCAGATGAAACAGAGGCTAAATGAGAATGAACG
AAAAAGAAAAAGGTGGTTATATTGGCAACCTATTCTCACTAAGATGGGGTTTATGTCAGTCATTTTGGTT
GGCGCTTTTGTTGGCTGGAGACTGTTTTTTTCAGCAAAATGCCCTATAAACAATTAATTTTGCCCAGCAAG
CTTCTGCACTAGTAACTGACAGTGCTACATTAATCATAGGGGTTTGTCTGCAGCAAACGCCTCATATCCC
AAAAACGGTGCAGTAGAATAGACATCAACCAGATAAGTGATATTTACAGTCACAAGCCCAACATCTCAGG
```
45
```
ACTCTTGACTGCAGGTTCCTCTGAACCCCAAACTGTAAATGGCTGTCTAAAATAAAGACATTCATGTTTG
TTAAAAACTGGTAAATTTTGCAACTGTATTCATACATGTCAAACACAGTATTTCACCTGACCAACATTGA
GATATCCTTTATCACAGGATTTGTTTTTGGAGGCTATCTGGATTTTAACCTGCACTTGATATAAGCAATA
AATATTGTGGTTTTATCTACGTTATTGGAAAGAAATGACATTTAAATAATGTGTGTAATGTATAATGTA
CTATTGACATGGGCATCAACACTTTTATTCTTAAGCATTTCAGGGTAAATATATTTTATAAGTATCTATT
```
50
```
TAATCTTTTGTAGTTAACTGTACTTTTTAAGAGCTCAATTTGAAAAATCTGTTACTAAAAAAAAAAATTG
TATGTCGATTGAATTGTACTGGATACATTTTCCATTTTTCTAAAAAGAAGTTTGATATGAGCAGTTAGAA
GTTGGAATAAGCAATTTCTACTATATATTGCATTTCTTTTATGTTTTACAGTTTTCCCCATTTTAAAAAG
AAAAGCAAACAAAGAAACAAAAGTTTTTCCTAAAAATATCTTTGAAGGAAATTCTCCTTACTGGGATAG
TCAGGTAAACAGTTGGTCAAGACTTTGTAAAGAAATTGGTTTCTGTAAATCCCATTATTGATATGTTTAT
```
55
```
TTTTCATGAAAATTTCAATGTAGTTGGGGTAGATTATGATTTAGGAAGCAAAGTAAGAAGCAGCATTTT
ATGATTCATAATTTCAGTTTACTAGACTGAAGTTTTGAAGTAAACCC
```

>gi|4506291|ref|NP_002819.1| protein tyrosine phosphatase, non-receptor type 2, isoform 1; T-cell protein tyrosine phosphatase [Homo sapiens] (SEQ ID NO:28)

60
```
MPTTIEREFEELDTQRRWQPLYLEIRNESHDYPHRVAKFPENRNRNRYRDVSPYDHSRVKLQNAENDYIN
```

ASLVDIEEAQRSYILTQGPLPNTCCHFWLMVWQQKTKAVVMLNRIVEKESVKCAQYWPTDDQEMLFKETG
FSVKLLSEDVKSYYTVHLLQLENINSGETRTISHFHYTTWPDFGVPESPASFLNFLFKVRESGSLNPDHG
PAVIHCSAGIGRSGTFSLVDTCLVLMEKGDDINIKQVLLNMRKYRMGLIQTPDQLRFSYMAIIEGAKCIK
GDSSIQKRWKELSKEDLSPAFDHSPNKIMTEKYNGNRIGLEEEKLTGDRCTGLSSKMQDTMEENSESALR
5   KRIREDRKATTAQKVQQMKQRLNENERKRKRWLYWQPILTKMGFMSVILVGAFVGWRLFFQQNAL

EDG1

10 >gi|22041681|ref|XM_001499.9| Homo sapiens endothelial differentiation,
    sphingolipid G-protein-coupled receptor, 1 (EDG1), mRNA (SEQ ID NO:29)
    TAAGTTTGCGAGAGCACTACGCAGTCAGTCGGGGGCAGCAGCAAGATGCGAAGCGAGCCGTACAGATCCC
    GGGCTCTCCGAACGCAACTTCGCCCTGCTTGAGCGAGGCTGCGGTTTCCGAGGCCCTCTCCAGCCAAGGA
    AAAGCTACACAAAAAGCCTGGATCACTCATCGAACCACCCCTGAAGCCAGTGAAGGCTCTCTCGCCTCGC
15 CCTCTAGCGTTCGTCTGGAGTAGCGCCACCCCGGCTTCCTGGGGACACAGGGTTGGCACC<u>ATG</u>GGGCCCA
    CCAGCGTCCCGCTGGTCAAGGCCCACCGCAGCTCGGTCTCTGACTACGTCAACTATGATATCATCGTCCG
    GCATTACAACTACACGGGAAAGCTGAATATCAGCGCGGACAAGGAGAACAGCATTAAACTGACCTCGGTG
    GTGTTCATTCTCATCTGCTGCTTTATCATCCTGGAGAACATCTTTGTCTTGCTGACCATTTGGAAAACCA
    AGAAATTCCACCGACCCATGTACTATTTTATTGGCAATCTGGCCCTCTCAGACCTGTTGGCAGGAGTAGC
20 CTACACAGCTAACCTGCTCTTGTCTGGGGCCACCACCTACAAGCTCACTCCCGCCCAGTGGTTTCTGCGG
    GAAGGGAGTATGTTTGTGGCCCTGTCAGCCTCCGTGTTCAGTCTCCTCGCCATCGCCATTGAGCGCTATA
    TCACAATGCTGAAAATGAAACTCCACAACGGGAGCAATAACTTCCGCCTCTTCCTGCTAATCAGCGCCTG
    CTGGGTCATCTCCCTCATCCTGGGTGGCCTGCCTATCATGGGCTGGAACTGCATCAGTGCGCTGTCCAGC
    TGCTCCACCGTGCTGCCGCTCTACCACAAGCACTATATCCTCTTCTGCACCACGGTCTTCACTCTGCTTC
25 TGCTCTCCATCGTCATTCTGTACTGCAGAATCTACTCCTTGGTCAG<u>GACTCGGAGCCGCCGCCTGACGTT
    CCGCAAGAACATTTCCAAGGCCAGCCGCAGCTCTGAGAAGTCGCTGGCGCTGCTCAAGACCGTAATTATC
    GTCCTGAGCGTCTTCATCGCCTGCTGGGCACCGCTCTTCATCCTGCTCCTGCTGGATGTGGGCTGCAAGG
    TGAAGACCTGTGACATCCTCTTCAGAGCGGAGTACTTCCTGGTGTTAGCTGTGCTCAACTCCGGCACCAA
    CCCCATCATTTACACTCTGACCAACAAGGAGATGCGTCGGGCCTTCATCCGGATCATGTCCTGCTGCAAG
30 TGCCCGAGCGGAGACTCTGCTGGCAAATTCAAGCGACCCATCATCGCCGGCATGGAATTCAGCCGCAGCA
    AATCGGACAATTCCTCCCACCCCCAGAAAGACGAAGGGGACAACCCAGAGACCATTATGTCTTCTGGAAA
    CGTCAACTCTTCTTCC<u>TAG</u>AACTGGAAGCTGTCCACCCACCGGAAGCGCTCTTTACTTGGTCGCTGGCCA
    CCCCAGTGTTTGGAAAAAAATCTCTGGGCTTCGACTGCTGCCAGGGAGGAGCTGCTGCAAGCCAGAGGGA
    GGAAGGGGGAGAATACGACAGCCTGGTGGTGTCGGGTGTTGGTGGGTAGAGTTAGTTCCTGTGAACAAT
35 GCACTGGGAAGGGTGGAGATCAGGTCCCGGCCTGGAATATATTTTCTACCCCCCTGGAGCTTTGATTTTG
    CACTGAGCCAAAGGTCTAGCATTGTCAAGCTCCTAAAGGGTTCATTTGGCCCCTCCTCAAAGACTAATGT
    CCCCATGTGAAAGCGTCTCTTTGTCTGGAGCTTTGAGGAGATGTTTTCCTTCACTTTAGTTTCAAACCCA
    AGTGAGTGTGTGCACTTCTGCTTCTTTAGGGATGCCCTGTACATCCCACACCCCACCCTCCCTTCCCTTC
    ATACCCCTCCTCAACGTTCTTTTACTTTATACTTTAACTACCTGAGAGTTATCAGAGCTGGGGTTGTGGA
40 ATGATCGATCATCTATAGCAAATAGGCTATGTTGAGTACGTAGGCTGTGGGAAGATGAAGATGGTTTGGA
    GGTGTAAAACAATGTCCTTCGCTGAGGCCAAAGTTTCCATGTAAGCGGGATCCGTTTTTTGGAATTTGGT
    TGAAGTCACTTTGATTTCTTTAAAAAACATCTTTTCAATGAAATGTGTTACCATTTCATATCCATTGAAG
    CCGAAATCTGCATAAGGAAGCCCACTTTATCTAAATGATATTAGCCAGGATCCTTGGTGTCCTAGGAGAA
    ACAGACAAGCAAACAAAGTGAAAACCGAATGGATTAACTTTTGCAAACCAAGGGAGATTTCTTAGCAAA
45 TGAGTCTAACAAATATGACATCTGTCTTTGGCACTTTTGTTGATGTTTATTTCAGAATGTTGTGTGATTC
    ATTTCAAGCAACAACATGGTTGTATTTTGTTGTGTTAAAAGTACTTTTCTTGATTTTTGAATGTATTTGT
    TTCAGCAGAAGTCATTTTATTGGATTTTTCTAACCCGTGTTAACACCATTGAATGTGTATTTCTTAAGAA
    AATACCACCCTCTTGTGCCCTTAAAAGCATTACTTTAACTGGTAGGGAACGCCAGAAACTTTTCAGTCCA
    GCTATTCATTAGATAGTAATTGAAGATATGTATAAATATTACAAAGAATAAAAATATATTACTGTCTCTT
50 TAGTATGGTTTTCAGTGCAATTAAACCGAGAGATGTCTTGTTTTTTTAAAAAGAATAGTATTTAATAGGT
    TTCTGACTTTTGTGGATCATTTTGCACATAGCTTTATCAACTTTTAAACATTAATAAACTGATTTTTTTA
    AAG

    >gi|11422839|ref|XP_001499.1| similar to endothelial differentiation,
55 sphingolipid G-protein-coupled receptor, 1; edg-1; sphingosine 1-phosphate
    receptor EDG1 [Homo sapiens] (SEQ ID NO:30)
    MGPTSVPLVKAHRSSVSDYVNYDIIVRHYNYTGKLNISADKENSIKLTSVVFILICCFIILENIFVLLTI
    WKTKKFHRPMYYFIGNLALSDLLAGVAYTANLLLSGATTYKLTPAQWFLREGSMFVALSASVFSLLAIAI
    ERYITMLKMKLHNGSNNFRLFLLISACWVISLILGGLPIMGWNCISALSSCSTVLPLYHKHYILFCTTVF
60 TLLLLSIVILYCRIYSLVRTRSRRLTFRKNISKASRSSEKSLALLKTVIIVLSVFIACWAPLFILLLLDV
    GCKVKTCDILFRAEYFLVLAVLNSGTNPIIYTLTNKEMRRAFIRIMSCCKCPSGDSAGKFKRPIIAGMEF

SRSKSDNSSHPQKDEGDNPETIMSSGNVNSSS

IL10Ra

>gi|4504632|ref|NM_001558.1| Homo sapiens interleukin 10 receptor, alpha
(IL10RA), mRNA (SEQ ID NO:31)
AAAGAGCTGGAGGCGCGCAGGCCGGCTCCGCTCCGGCCCCGGACGATGCGGCGCGCCCAGG<u>ATG</u>CTGCCG
TGCCTCGTAGTGCTGCTGGCGGCGCTCCTCAGCCTCCGTCTTGGCTCAGACGCTCATGGGACAGAGCTGC
CCAGCCCTCCGTCTGTGTGGTTTGAAGCAGAATTTTTCCACCACATCCTCCACTGGACACCCATCCCAAA
TCAGTCTGAAAGTACCTGCTATGAAGTGGCGCTCCTGAGGTATGGAATAGAGTCCTGGAACTCCATCTCC
AACTGTAGCCAGACCCTGTCCTATGACCTTACCGCAGTGACCTTGGACCTGTACCACAGCAATGGCTACC
GGGCCAGAGTGCGGGCTGTGGACGGCAGCCGGCACTCCAACTGGACCGTCACCAACACCCGCTTCTCTGT
GGATGAAGTGACTCTGACAGTTGGCAGTGTGAACCTAGAGATCCACAATGGCTTCATCCTCGGGAAGATT
CAGCTACCCAGGCCCAAGATGGCCCCCGCGAATGACACATATGAAAGCATCTTCAGTCACTTCCGAGAGT
ATGAGATTGCCATTCGCAAGGTGCCGGGAAACTTCACGTTCACACACAAGAAAGTAAAACATGAAAACTT
CAGCCTCCTAACCTCTGGAGAAGTGGGAGAGTTCTGTGTCCAGGTGAAACCATCTGTCGCTTCCCGAAGT
AACAAGGGGATGTGGTCTAAAGAGGAGTGCATCTCCCTCACCAGGCAGTATTTCACCGTGACCAACGTCA
TCATCTTCTTTGCCTTTGTCCTGCTGCTCTCCGGAGCCCTCGCCTACTGCCTGGCCCTCCAGCTGTATGT
GCGGCGCCGAAAGAAGCTACCCAGTGTCCTGCTCTTCAAGAAGCCCAGCCCCTTCATCTTCATCAGCCAG
CGTCCCTCCCCAGAGACCCAAGACACCATCCACCCGCTTGATGAGGAGGCCTTTTTGAAGGTGTCCCCAG
AGCTGAAGAACTTGGACCTGCACGGCAGCACAGACAGTGGCTTTGGCAGCACCAAGCCATCCCTGCAGAC
TGAAGAGCCCCAGTTCCTCCTCCCTGACCCTCACCCCCAGGCTGACAGAACGCTGGGAAACGGGGAGCCC
CCTGTGCTGGGGGACAGCTGCAGTAGTGGCAGCAGCAATAGCACAGACAGCGGGATCTGCCTGCAGGAGC
CCAGCCTGAGCCCCAGCACAGGGCCCACCTGGGAGCAACAGGTGGGGAGCAACAGCAGGGGCCAGGATGA
CAGTGGCATTGACTTAGTTCAAAACTCTGAGGGCCGGGCTGGGGACACACAGGGTGGCTCGGCCTTGGGC
CACCACAGTCCCCCGGAGCCTGAGGTGCCTGGGGAAGAAGACCCAGCTGCTGTGGCATTCCAGGGTTACC
TGAGGCAGACCAGATGTGCTGAAGAGAAGGCAACCAAGACAGGCTGCCTGGAGGAAGAATCGCCCTTGAC
AGATGGCCTTGGCCCCAAATTCGGGAGATGCCTGGTTGATGAGGCAGGCTTGCATCCACCAGCCCTGGCC
AAGGGCTATTTGAAACAGGATCCTCTAGAAATGACTCTGGCTTCCTCAGGGGCCCCAACGGGACAGTGGA
ACCAGCCCACTGAGGAATGGTCACTCCTGGCCTTGAGCAGCTGCAGTGACCTGGGAATATCTGACTGGAG
CTTTGCCCATGACCTTGCCCCTCTAGGCTGTGTGGCAGCCCCAGGTGGTCTCCTGGGCAGCTTTAACTCA
GACCTGGTCACCCTGCCCCTCATCTCTAGCCTGCAGTCAAGTGAGT<u>GA</u>CTCGGGCTGAGAGGCTGCTTTT
GATTTTAGCCATGCCTGCTCCTCTGCCTGGACCAGGAGGAGGGCCCTGGGGCAGAAGTTAGGCACGAGGC
AGTCTGGGCACTTTTCTGCAAGTCCACTGGGGCTGGCCCAGCCAGGCTGCAGGGCTGGTCAGGGTGTCTG
GGGCAGGAGGAGGCCAACTCACTGAACTAGTGCAGGGTATGTGGGTGGCACTGACCTGTTCTGTTGACTG
GGGCCCTGCAGACTCTGGCAGAGCTGAGAAGGGCAGGGACCTTCTCCCTCCTAGGAACTCTTTCCTGTAT
CATAAAGGATTATTTGCTCAGGGGAACCATGGGGCTTTCTGGAGTTGTGGTGAGGCCACCAGGCTGAAGT
CAGCTCAGACCCAGACCTCCCTGCTTAGGCCACTCGAGCATCAGAGCTTCCAGCAGGAGGAAGGGCTGTA
GGAATGGAAGCTTCAGGGCCTTGCTGCTGGGGTCATTTTTAGGGGAAAAAGGAGGATATGATGGTCACAT
GGGGAACCTCCCCTCATCGGGCCTCTGGGGCAGGAAGCTTGTCACTGGAAGATCTTAAGGTATATATTTT
CTGGACACTCAAACACATCATAATGGATTCACTGAGGGGAGACAAAGGGAGCCGAGACCCTGGATGGGGC
TTCCAGCTCAGAACCCATCCCTCTGGTGGGTACCTCTGGCACCCATCTGCAAATATCTCCCTCTCTCCAA
CAAATGGAGTAGCATCCCCCTGGGGCACTTGCTGAGGCCAAGCCACTCACATCCTCACTTTGCTGCCCCA
CCATCTTGCTGACAACTTCCAGAGAAGCCATGGTTTTTTGTATTGGTCATAACTCAGCCCTTTGGGCGGC
CTCTGGGCTTGGGCACCAGCTCATGCCAGCCCCAGAGGGTCAGGGTTGGAGGCCTGTGCTTGTGTTTGCT
GCTAATGTCCAGCTACAGACCCAGAGGATAAGCCACTGGGCACTGGCTGGGGTCCCTGCCTTGTTGGTG
TTCAGCTGTGTGATTTTGGACTAGCCACTTGTCAGAGGGCCTCAATCTCCCATCTGTGAAATAAGGACTC
CACCTTTAGGGGACCCTCCATGTTTGCTGGGTATTAGCCAAGCTGGTCCTGGGAGAATGCAGATACTGTC
CGTGGACTACCAAGCTGGCTTGTTTCTTATGCCAGAGGCTAACAGATCCAATGGGAGTCCATGGTGTCAT
GCCAAGACAGTATCAGACACAGCCCCAGAAGGGGGCATTATGGGCCCTGCCTCCCCATAGGCCATTTGGA
CTCTGCCTTCAAACAAAGGCAGTTCAGTCCACAGGCATGGAAGCTGTGAGGGGACAGGCCTGTGCGTGCC
ATCCAGAGTCATCTCAGCCCTGCCTTTCTCTGGAGCATTCTGAAAACAGATATTCTGGCCCAGGGAATCC
AGCCATGACCCCCACCCCTCTGCCAAAGTACTCTTAGGTGCCAGTCTGGTAACTGAACTCCCTCTGGAGG
CAGGCTTGAGGGAGGATTCCTCAGGGTTCCCTTGAAAGCTTTATTTATTTATTTTGTTCATTTATTTATT
GGAGAGGCAGCATTGCACAGTGAAAGAATTCTGGATATCTCAGGAGCCCCGAAATTCTAGCTCTGACTTT
GCTGTTTCCAGTGGTATGACCTTGGAGAAGTCACTTATCCTCTTGGAGCCTCAGTTTCCTCATCTGCAGA
ATAATGACTGACTTGTCTAATTCATAGGGATGTGAGGTTCTGCTGAGGAAATGGGTATGAATGTGCCTTG
AACACAAAGCTCTGTCAATAAGTGATACATGTTTTTTATTCCAATAAATTGTCAAGACCACA

>gi|4504633|ref|NP_001549.1| interleukin 10 receptor, alpha; Interleukin-10
receptor [Homo sapiens] (SEQ ID NO:32)
MLPCLVVLLAALLSLRLGSDAHGTELPSPPSVWFEAEFFHHILHWTPIPNQSESTCYEVALLRYGIESWN
SISNCSQTLSYDLTAVTLDLYHSNGYRARVRAVDGSRHSNWTVTNTRFSVDEVTLTVGSVNLEIHNGFIL

5   GKIQLPRPKMAPANDTYESIFSHFREYEIAIRKVPGNFTFTHKKVKHENFSLLTSGEVGEFCVQVKPSVA
SRSNKGMWSKEECISLTRQYFTVTNVIIFFAFVLLLSGALAYCLALQLYVRRRKKLPSVLLFKKPSPFIF
ISQRPSPETQDTIHPLDEEAFLKVSPELKNLDLHGSTDSGFGSTKPSLQTEEPQFLLPDPHPQADRTLGN
GEPPVLGDSCSSGSSNSTDSGICLQEPSLSPSTGPTWEQQVGSNSRGQDDSGIDLVQNSEGRAGDTQGGS
ALGHHSPPEPEVPGEEDPAAVAFQGYLRQTRCAEEKATKTGCLEEESPLTDGLGPKFGRCLVDEAGLHPP

10   ALAKGYLKQDPLEMTLASSGAPTGQWNQPTEEWSLLALSSCSDLGISDWSFAHDLAPLGCVAAPGGLLGS
FNSDLVTLPLISSLQSSE

integrin a2

15

>gi|6006008|ref|NM_002203.2| Homo sapiens integrin, alpha 2 (CD49B, alpha 2
subunit of VLA-2 receptor) (ITGA2), mRNA (SEQ ID NO:33)
CTGCAAACCCAGCGCAACTACGGTCCCCCGGTCAGACCCAGGATGGGGCCAGAACGGACAGGGGCCGCGC
CGCTGCCGCTGCTGCTGGTGTTAGCGCTCAGTCAAGGCATTTTAAATTGTTGTTTGGCCTACAATGTTGG

20   TCTCCCAGAAGCAAAAATATTTTCCGGTCCTTCAAGTGAACAGTTTGGGTATGCAGTGCAGCAGTTTATA
AATCCAAAAGGCAACTGGTTACTGGTTGGTTCACCCTGGAGTGGCTTTCCTGAGAACCGAATGGGAGATG
TGTATAAATGTCCTGTTGACCTATCCACTGCCACATGTGAAAAACTAAATTTGCAAACTTCAACAAGCAT
TCCAAATGTTACTGAGATGAAAACCAACATGAGCCTCGGCTTGATCCTCACCAGGAACATGGGAACTGGA
GGTTTTCTCACATGTGGTCCTCTGTGGGCACAGCAATGTGGGAATCAGTATTACACAACGGGTGTGTGTT

25   CTGACATCAGTCCTGATTTTCAGCTCTCAGCCAGCTTCTCACCTGCAACTCAGCCCTGCCCTTCCCTCAT
AGATGTTGTGGTTGTGTGTGATGAATCAAATAGTATTTATCCTTGGGATGCAGTAAAGAATTTTTTGGAA
AAATTTGTACAAGGCCTTGATATAGGCCCCACAAAGACACAGGTGGGGTTAATTCAGTATGCCAATAATC
CAAGAGTTGTGTTTAACTTGAACACATATAAAACCAAAGAAGAAATGATTGTAGCAACATCCCAGACATC
CCAATATGGTGGGGACCTCACAAACACATTCGGAGCAATTCAATATGCAAGAAAATATGCCTATTCAGCA

30   GCTTCTGGTGGGCGACGAAGTGCTACGAAAGTAATGGTAGTTGTAACTGACGGTGAATCACATGATGGTT
CAATGTTGAAAGCTGTGATTGATCAATGCAACCATGACAATATACTGAGGTTTGGCATAGCAGTTCTTGG
GTACTTAAACAGAAACGCCCTTGATACTAAAAATTTAATAAAAGAAATAAAAGCGATCGCTAGTATTCCA
ACAGAAAGATACTTTTTCAATGTGTCTGATGAAGCAGCTCTACTAGAAAAGGCTGGGACATTAGGAGAAC
AAATTTTCAGCATTGAAGGTACTGTTCAAGGAGGAGACAACTTTCAGATGGAAATGTCACAAGTGGGATT

35   CAGTGCAGATTACTCTTCTCAAAATGATATTCTGATGCTGGGTGCAGTGGGAGCTTTTGGCTGGAGTGGG
ACCATTGTCCAGAAGACATCTCATGGCCATTTGATCTTTCCTAAACAAGCCTTTGACCAAATTCTGCAGG
ACAGAAATCACAGTTCATATTTAGGTTACTCTGTGGCTGCAATTTCTACTGGAGAAAGCACTCACTTTGT
TGCTGGTGCTCCTCGGGCAAATTATACCGGCCAGATAGTGCTATATAGTGTGAATGAGAATGGCAATATC
ACGGTTATTCAGGCTCACCGAGGTGACCAGATTGGCTCCTATTTTGGTAGTGTGCTGTGTTCAGTTGATG

40   TGGATAAAGACACCATTACAGACGTGCTCTTGGTAGGTGCACCAATGTACATGAGTGACCTAAAGAAAGA
GGAAGGAAGAGTCTACCTGTTTACTATCAAAAAGGGCATTTTGGGTCAGCACCAATTTCTTGAAGGCCCC
GAGGGCATTGAAAACACTCGATTTGGTTCAGCAATTGCAGCTCTTTCAGACATCAACATGGATGGCTTTA
ATGATGTGATTGTTGGTTCACCACTAGAAAATCAGAATTCTGGAGCTGTATACATTTACAATGGTCATCA
GGGCACTATCCGCACAAAGTATTCCCAGAAAATCTTGGGATCCGATGGAGCCTTTAGGAGCCATCTCCAG

45   TACTTTGGGAGGTCCTTGGATGGCTATGGAGATTTAAATGGGGATTCCATCACCGATGTGTCTATTGGTG
CCTTTGGACAAGTGGTTCAACTCTGGTCACAAAGTATTGCTGATGTAGCTATAGAAGCTTCATTCACACC
AGAAAAAATCACTTTGGTCAACAAGAATGCTCAGATAATTCTCAAACTCTGCTTCAGTGCAAAGTTCAGA
CCTACTAAGCAAAACAATCAAGTGGCCATTGTATATAACATCACACTTGATGCAGATGGATTTTCATCCA
GAGTAACCTCCAGGGGGTTATTTAAAGAAAACAATGAAAGGTGCCTGCAGAAGAATATGGTAGTAAATCA

50   AGCACAGAGTTGCCCCGAGCACATCATTTATATACAGGAGCCCTCTGATGTTGTCAACTCTTTGGATTTG
CGTGTGGACATCAGTCTGGAAAACCCTGGCACTAGCCCTGCCCTTGAAGCCTATTCTGAGACTGCCAAGG
TCTTCAGTATTCCTTTCCACAAAGACTGTGGTGAGGATGGACTTTGCATTTCTGATCTAGTCCTAGATGT
CCGACAAATACCAGCTGCTCAAGAACAACCCTTTATTGTCAGCAACCAAAACAAAGGTTAACATTTTCA
GTAACACTGAAAAATAAAAGGGAAAGTGCATACAACACTGGAATTGTTGTTGATTTTTCAGAAACTTGT

55   TTTTTGCATCATTCTCCCTACCGGTTGATGGGACAGAAGTAACATGCCAGGTGGCTGCATCTCAGAAGTC
TGTTGCCTGCGATGTAGGCTACCCTGCTTTAAAGAGAGAACAACAGGTGACTTTTACTATTAACTTTGAC
TTCAATCTTCAAAACCTTCAGAATCAGGCGTCTCTCAGTTTCCAAGCCTTAAGTGAAAGCCAAGAAGAAA
ACAAGGCTGATAATTTGGTCAACCTCAAAATTCCTCTCCTGTATGATGCTGAAATTCACTTAACAAGATC
TACCAACATAAATTTTTATGAAATCTCTTCGGATGGGAATGTTCCTTCAATCGTGCACAGTTTTGAAGAT

60   GTTGGTCCAAAATTCATCTTCTCCCTGAAGGTAACAACAGGAAGTGTTCCAGTAAGCATGGCAACTGTAA
TCATCCACATCCCTCAGTATACCAAAGAAAAGAACCCACTGATGTACCTAACTGGGGTGCAAACAGACAA

77

```
          GGCTGGTGACATCAGTTGTAATGCAGATATCAATCCACTGAAAATAGGACAAACATCTTCTTCTGTATCT
          TTCAAAAGTGAAAATTTCAGGCACACCAAAGAATTGAACTGCAGAACTGCTTCCTGTAGTAATGTTACCT
          GCTGGTTGAAAGACGTTCACATGAAAGGAGAATACTTTGTTAATGTGACTACCAGAATTTGGAACGGGAC
          TTTCGCATCATCAACGTTCCAGACAGTACAGCTAACGGCAGCTGCAGAAATCAACACCTATAACCCTGAG
    5     ATATATGTGATTGAAGATAACACTGTTACGATTCCCCTGATGATAATGAAACCTGATGAGAAAGCCGAAG
          TACCAACAGGAGTTATAATAGGAAGTATAATTGCTGGAATCCTTTTGCTGTTAGCTCTGGTTGCAATTTT
          ATGGAAGCTCGGCTTCTTCAAAAGAAAATATGAAAGATGACCAAAAATCCAGATGAGATTGATGAGACC
          ACAGAGCTCAGTAGCTGAACCAGCAGACCTACCTGCAGTGGGAACCGGCAGCATCCCAGCCAGGGTTTGC
          TGTTTGCGTGCATGGATTTCTTTTTAAATCCCATATTTTTTTATCATGTCGTAGGTAAACTAACCTGGT
   10     ATTTTAAGAGAAAACTGCAGGTCAGTTTGGATGAAGAAATTGTGGGGGGTGGGGGAGGTGCGGGGGGCAG
          GTAGGGAAATAATAGGGAAAATACCTATTTTATATGATGGGGGAAAAAAAGTAATCTTTAAACTGGCTGG
          CCCAGAGTTTACATTCTAATTTGCATTGTGTCAGAAACATGAAATGCTTCCAAGCATGACAACTTTTAAA
          GAAAAATATGATACTCTCAGATTTTAAGGGGGAAAACTGTTCTCTTTAAAATATTTGTCTTTAAACAGCA
          ACTACAGAAGTGGAAGTGCTTGATATGTAAGTACTTCCACTTGTGTATATTTTAATGAATATTGATGTTA
   15     ACAAGAGGGGAAAACAAAACACAGGTTTTTTCAATTTATGCTGCTCATCCAAAGTTGCCACAGATGATAC
          TTCCAAGTGATAATTTTATTTATAAACTAGGTAAAATTTGTTGTTGGTTCCTTTTATACCACGGCTGCCC
          CTTCCACACCCCATCTTGCTCTAATGATCAAAACATGCTTGAATAACTGAGCTTAGAGTATACCTCCTAT
          ATGTCCATTTAAGTTAGGAGAGGGGGCGATATAGAGACTAAGGCACAAAATTTTGTTTAAAACTCAGAAT
          ATAACATTTATGTAAAATCCCATCTGCTAGAAGCCCATCCTGTGCCAGAGGAAGGAAAAGGAGGAAATTT
   20     CCTTTCTCTTTTAGGAGGCACAACAGTTCTCTTCTAGGATTTGTTTGGCTGACTGGCAGTAACCTAGTGA
          ATTTTTGAAAGATGAGTAATTTCTTTGGCAACCTTCCTCCTCCCTTACTGAACCACTCTCCCACCTCCTG
          GTGGTACCATTATTATAGAAGCCCTCTACAGCCTGACTTTCTCTCCAGCGGTCCAAAGTTATCCCCTCCT
          TTACCCCTCATCCAAAGTTCCCACTCCTTCAGGACAGCTGCTGTGCATTAGATATTAGGGGGGAAAGTCA
          TCTGTTTAATTTACACACTTGCATGAATTACTGTATATAAACTCCTTAACTTCAGGGAGCTATTTTCATT
   25     TAGTGCTAAACAAGTAAGAAAAATAAGCTAGAGTGAATTTCTAAATGTTGGAATGTTATGGGATGTAAAC
          AATGTAAAGTAAAACACTCTCAGGATTTCACCAGAAGTTACAGATGAGGCACTGGAAACCACCACCAAAT
          TAGCAGGTGCACCTTCTGTGGCTGTCTTGTTTCTGAAGTACTTTTTCTTCCACAAGAGTGAATTTGACCT
          AGGCAAGTTTGTTCAAAAGGTAGATCCTGAGATGATTGGTCAGATTGGGATAAGGCCCAGCAATCTGCA
          TTTTAACAAGCACCCCAGTCACTAGGATGCAGATGGACCACACTTTGAGAAACACCACCCATTTCTACTT
   30     TTTGCACCTTATTTTCTCTGTTCCTGAGCCCCCACATTCTCTAGGAGAAACTTAGATTAAAATTCACAGA
          CACTACATATCTAAAGCTTTGACAAGTCCTTGACCTCTATAAACTTCAGAGTCCTCATTATAAATGGGA
          AGACTGAGCTGGAGTTCAGCAGTGATGCTTTTTAGTTTTAAAAGTCTATGATCTGATCTGGACTTCCTAT
          AATACAAATACACAATCCTCCAAGAATTTGACTTGGAAAAG
```

   35   >gi|4504743|ref|NP_002194.1| integrin alpha 2 precursor; Integrin, alpha-2
          (CD49B; alpha-2 subunit of VLA-2 receptor; platelet antigen Br) [Homo
          sapiens] (SEQ ID NO:34)
          MGPERTGAAPLPLLLVLALSQGILNCCLAYNVGLPEAKIFSGPSSEQFGYAVQQFINPKGNWLLVGSPWS
          GFPENRMGDVYKCPVDLSTATCEKLNLQTSTSIPNVTEMKTNMSLGLILTRNMGTGGFLTCGPLWAQQCG
   40   NQYYTTGVCSDISPDFQLSASFSPATQPCPSLIDVVVVCDESNSIYPWDAVKNFLEKFVQGLDIGPTKTQ
          VGLIQYANNPRVVFNLNTYKTKEEMIVATSQTSQYGGDLTNTFGAIQYARKYAYSAASGGRRSATKVMVV
          VTDGESHDGSMLKAVIDQCNHDNILRFGIAVLGYLNRNALDTKNLIKEIKAIASIPTERYFFNVSDEAAL
          LEKAGTLGEQIFSIEGTVQGGDNFQMEMSQVGFSADYSSQNDILMLGAVGAFGWSGTIVQKTSHGHLIFP
          KQAFDQILQDRNHSSYLGYSVAAISTGESTHFVAGAPRANYTGQIVLYSVNENGNITVIQAHRGDQIGSY
   45   FGSVLCSVDVDKDTITDVLLVGAPMYMSDLKKEEGRVYLFTIKKGILGQHQFLEGPEGIENTRFGSAIAA
          LSDINMDGFNDVIVGSPLENQNSGAVYIYNGHQGTIRTKYSQKILGSDGAFRSHLQYFGRSLDGYGDLNG
          DSITDVSIGAFGQVVQLWSQSIADVAIEASFTPEKITLVNKNAQIILKLCFSAKFRPTKQNNQVAIVYNI
          TLDADGFSSRVTSRGLFKENNERCLQKNMVVNQAQSCPEHIIYIQEPSDVVNSLDLRVDISLENPGTSPA
          LEAYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKNKRESAYNTG
   50   IVVDFSENLFFASFSLPVDGTEVTCQVAASQKSVACDVGYPALKREQQVTFTINFDFNLQNLQNQASLSF
          QALSESQEENKADNLVNLKIPLLYDAEIHLTRSTNINFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTTG
          SVPVSMATVIIHIPQYTKEKNPLMYLTGVQTDKAGDISCNADINPLKIGQTSSSVSFKSENFRHTKELNC
          RTASCSNVTCWLKDVHMKGEYFVNVTTRIWNGTFASSTFQTVQLTAAAEINTYNPEIYVIEDNTVTIPLM
          IMKPDEKAEVPTGVIIGSIIAGILLLLALVAILWKLGFFKRKYEKMTKNPDEIDETTELSS

   55

          Enolase 1a

          >gi|4503570|ref|NM_001428.1| Homo sapiens enolase 1, (alpha) (ENO1), mRNA
   60   (SEQ ID NO:35)
          ACGGAGATCTCGCCGGCTTTACGTTCACCTCGGTGTCTGCAGCACCCTCCGCTTCCTCTCCTAGGCGACG

AGACCCAGTGGCTAGAAGTTCACC<u>ATG</u>TCTATTCTCAAGATCCATGCCAGGGAGATCTTTGACTCTCGCG
GGAATCCCACTGTTGAGGTTGATCTCTTCACCTCAAAAGGTCTCTTCAGAGCTGCTGTGCCCAGTGGTGC
TTCAACTGGTATCTATGAGGCCCTAGAGCTCCGGGACAATGATAAGACTCGCTATATGGGGAAGGGTGTC
TCAAAGGCTGTTGAGCACATCAATAAAACTATTGCGCCTGCCCTGGTTAGCAAGAAACTGAACGTCACAG
AACAAGAGAAGATTGACAAACTGATGATCGAGATGGATGGAACAGAAAATAAATCTAAGTTTGGTGCGAA
CGCCATTCTGGGGGTGTCCCTTGCCGTCTGCAAAGCTGGTGCCGTTGAGAAGGGGGTCCCCCTGTACCGC
CACATCGCTGACTTGGCTGGCAACTCTGAAGTCATCCTGCCAGTCCCGGCGTTCAATGTCATCAATGGCG
GTTCTCATGCTGGCAACAAGCTGGCCATGCAGGAGTTCATGATCCTCCCAGTCGGTGCAGCAAACTTCAG
GGAAGCCATGCGCATTGGAGCAGAGGTTTACCACAACCTGAAGAATGTCATCAAGGAGAAATATGGGAAA
GATGCCACCAATGTGGGGGATGAAGGCGGGTTTGCTCCCAACATCCTGGAGAATAAAGAAGGCCTGGAGC
TGCTGAAGACTGCTATTGGGAAAGCTGGCTACACTGATAAGGTGGTCATCGGCATGGACGTAGCGGCCTC
CGAGTTCTTCAGGTCTGGGAAGTATGACCTGGACTTCAAGTCTCCCGATGACCCCAGCAGGTACATCTCG
CCTGACCAGCTGGCTGACCTGTACAAGTCCTTCATCAAGGACTACCCAGTGGTGTCTATCGAAGATCCCT
TTGACCAGGATGACTGGGGAGCTTGGCAGAAGTTCACAGCCAGTGCAGGAATCCAGGTAGTGGGGGATGA
TCTCACAGTGACCAACCCAAAGAGGATCGCCAAGGCCGTGAACGAGAAGTCCTGCAACTGCCTCCTGCTC
AAAGTCAACCAGATTGGCTCCGTGACCGAGTCTCTTCAGGCGTGCAAGCTGGCCCAGGCCAATGGTTGGG
GCGTCATGGTGTCTCATCGTTCGGGGGAGACTGAAGATACCTTCATCGCTGACCTGGTTGTGGGGCTGTG
CACTGGGCAGATCAAGACTGGTGCCCCTTGCCGATCTGAGCGCTTGGCCAAGTACAACCAGCTCCTCAGA
ATTGAAGAGGAGCTGGGCAGCAAGGCTAAGTTTGCCGGCAGGAACTTCAGAAACCCCTTGGCCAAGT<u>AAG</u>
CTGTGGGCAGGCAAGCCTTCGGTCACCTGTTGGCTACACAGACCCCTCCCCTCGTGTCAGCTCAGGCAGC
TCGAGGCCCCCGACCAACACTTGCAGGGGTCCCTGCTAGTTAGCGCCCCACCGCCGTGGAGTTCGTACCG
CTTCCTTAGAACTTCTACAGAAGCCAAGCTCCCTGGAGCCCTGTTGGCAGCTCTAGCTTTTGCAGTCGTG
TAATGGGCCCAAGTCATTGTTTTTCTCGCCTCACTTTCCACCAAGTGTCTAGAGTCATGTGAGCCTCGTG
TCATCTCCGGGGTGGCCACAGGCTAGATCCCCGGTGGTTTTGTGCTCAAAATAAAAAGCCTCAGTGACCC
ATGAG

>gi|4503571|ref|NP_001419.1| enolase 1; phosphopyruvate hydratase; enolase-
1, alpha; enolase 1, (alpha)-like 1; MYC promoter-binding protein 1; non-
neural enolase; 2-phospho-D-glycerate hydro-lyase; crystallin, tau,
included [Homo sapiens] (SEQ ID NO:36)
MSILKIHAREIFDSRGNPTVEVDLFTSKGLFRAAVPSGASTGIYEALELRDNDKTRYMGKGVSKAVEHIN
KTIAPALVSKKLNVTEQEKIDKLMIEMDGTENKSKFGANAILGVSLAVCKAGAVEKGVPLYRHIADLAGN
SEVILPVPAFNVINGGSHAGNKLAMQEFMILPVGAANFREAMRIGAEVYHNLKNVIKEKYGKDATNVGDE
GGFAPNILENKEGLELLKTAIGKAGYTDKVVIGMDVAASEFFRSGKYDLDFKSPDDPSRYISPDQLADLY
KSFIKDYPVVSIEDPFDQDDWGAWQKFTASAGIQVVGDDLTVTNPKRIAKAVNEKSCNCLLLKVNQIGSV
TESLQACKLAQANGWGVMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKYNQLLRIEEELGSK
AKFAGRNFRNPLAK

PRSM1

>gi|1354930|gb|U58048.1|HSU58048 Human metallopeptidase PRSM1 mRNA,
complete cds (SEQ ID NO:37)
AGCGACTCACAAAAGGTCCTCGGCCACGGCGTGCGTCACC<u>ATG</u>GCGACCGCCGCCGGCAGCCGCGCCCCG
CCCCTCTGGCGGGACCGGCCACCATCCTCTCGCGAGGAGCATCCCGTGCGACCGGAAGTGGGGCGGCGAC
CCCGGAAGTCCCCGCCGGGTGCAGCTTGGTCGGTTCGATCGCCGCCGGGACCTGACACCGCCCGGAGTTG
GCGTCCCTTCTCCCTCTCCGAGTGCTGCTCCTGTCATTGTGGCCATGGACGATACCCTGTTCCAGTTGAA
GTTCACGGCGAAGCAGCTGGAGAAGCTGGCCAAGAAGGCGGAGAAGGACTCCAAGGCGGAGCAGGCCAAA
GTGAAGAAGGCCCTTCTGCAGAAAAATGTAGAGTGTGCCCGTGTGTATGCCGAGAACGCCATCCGCAAGA
AGAACGAAGGTGTGAACTGGCTTCGGATGGCGTCCCGCGTAGACGCAGTGGCCTCCAAGGTGGACACAGC
TGTGACTATGAAGGGGGTGACCAAGAATATGGCCCAGGTGACCAAAGCCCTGGACAAGGCCCTGAGCACC
ATGGACCTGCAGAAGGTCTCCTCAGTGATGGACAGGTTCGAGCAGCAGGTGCAGAACCTGGACGTCCATA
CATCGGTGATGGAGGACTCCATGAGCTCGGCCACCACCCTGACCACGCCGCAGGAGCAGGTGGACAGCCT
CATCATGCAGATCGCCGAGGAGAATGGCCTGGAGGTGCTGGACCAGCTCAGCCAGCTGCCCGAGGGCGCC
TCTGCCGTGGGCGAGAGCTCTGTGCGCAGCCAGGAGGACCAGCTGTCACGGAGGTTGGCCGCCTTGAGGA
ACTAGCCGTGCCCCGCCGGTGTGCACCGCCTCTGCCCCGTGATGTGCTGGAAGGCTCCTGTCCTCTCCCC
ACCGCGTCTTGCCTTTGTGCTGACCCCGCGGGGCTGCGGCCGGCAGCCACTCTGCGTCTCTCACCTGCCA
GGCCTGCGTGGCCTT<u>TAG</u>GGTTGTTCCTGTTCTTTTAGGTTGGGCGGTGGGTCTGTGTCCTGGTGTTGAGT
TTCTGCAAATTTCTGGGGGTGATTTCTGTGACTCTGGGCCCACAGCGGGGAGGCCAAGAGGGGCCCTGTG
GACTTTCACCCAGCACTGTGGGGGCCTTCAGACTCTGGGGCAGCAGACATGCTGCTTCCCATCAGCCAGA
GGGGGTCAGGGCTGCCCTGTTGCCAAACAACTCCCTGAGGCCTCTCCGCACCACCTCAGCGGGCAGGAGG

TCCCACCATGTGGACAGACATAGCCCAAGGAGGCACCACAGGTCTATGTGTGCTGGGGGATGTCAGGTGC
CACCCAACGCTGTCCTGGTGGTATTTACAATGACATCCTCCTCCTCCATCACTCCAGGGGTGGTGTCTCG
GCCGCCCCTACCAGCTGGCTGAGCCCCCTGGCCTCCTGCGCTCCCTCACTTCCCTCAGTTCCCAAAGCTG
CCCAGTCCATGGGGACAGAACCGTCACTCAGATCCACATTCAAGTGTGCCCACCCTGCAGTCTTCATCCT
5   CACTCAGCTGCTGCCTCTGGAGGTGCCTTTGGCCACATGTGCTGTGCTGTTTGTCTCCTCGACAGGGAGC
CTGTCCACCAGCAGGCTGCGGTCCCAGCGGGTGCGTCTGCAGCTCCTCCCCTTGGGCAGGCTGGTTCTCC
CGGAGGACCTTTCCTTGGGGCCCTGCTTCATGACGATGCTGCCTGTGTCACCCTCTACCATCTGTAAACA
ACTGGGTGCCTTCCCCGACCACACCCCAATGCCTTCCCAGCTTGGAAGCCAAGGCAGCTGATGAAGGGAG
CTCAGGAGAGCCGTCTTCAGCTGGGCTGGGGTTGGGGCTGCTGTGAGGAAAACCTGCCATTGTGGCCCTG
10   GAGAGTCACCAGCAGCTCTTGGGAAGGACTTGCTGGGAGGCTGAGAGAGGCTTTGGGCACAGCCTGCTGT
CTTTTCCATTTCCTAAAGTTTACTTCATTGTCTTGAGGCTTCCAGGTTTTGTTTTTGTTTTTGCCAAAGT
AGAAAAGGCAGGTGGTGGGCGGCTGGCAGGGAGTGCGGGTCCCCGCCCCTCTTCAGTCCTGCCCTCCCCT
CCTCAGTCCTGCCCACCCCGTGCAGCCCATGCTGAGGCTGCAGTGGTGTCGTGGGTGTTACGTGCAGGAA
CGTGGAGACCCTGACGTGGGCTCACTGCATTTGGTTTTCTTTTCAGAACTTGGGAGCCCCCAGGGAGGGG
15   CTAGTGTTGGTAGGTCCTAGACGTGGTTCCCTCCAGCCTCCCCAAAATCAACCCTGGTGTTGAGAGAACG
TCCTTCTGTCCATCGTGGGTAACAGCCTTGGGGAGGGTGCAGAGCTCTGCAGAGCCATGGGCCAGGTGGG
GCTGCCTCAGTCCTGTCCCCTTGGGCACTGAGGAGAGGGGCCCATTCACCTTTCTCCTAGAATGCTGTTG
TAAATAAACAAATGGATCCCTGGA

20   >gi|1354931|gb|AAC50775.1| PRSM1 (SEQ ID NO:38)
MATAAGSRAPPLWRDRPPSSREEHPVRPEVGRRPRKSPPGAAWSVRSPPGPDTARSWRPFSLSECCSCHC
GHGRYPVPVEVHGEAAGEAGQEGGEGLQGGAGQSEEGPSAEKCRVCPCVCRERHPQEERRCELASDGVPR
RRSGLQGGHSCDYEGGDQEYGPGDQSPGQGPEHHGPAEGLLSDGQVRAAGAEPGRPYIGDGGLHELGHHP
DHAAGAGGQPHHADRRGEWPGGAGPAQPAARGRLCRGRELCAQPGGPAVTEVGRLEELAVPRRCAPPLPR
25   DVLEGSCPLPTASCLCADPAGLRPAATLRLSPARPAWP

CLN2

30   >gi|5597012|ref|NM_000391.2| Homo sapiens ceroid-lipofuscinosis, neuronal
2, late infantile (Jansky-Bielschowsky disease) (CLN2), mRNA (SEQ ID NO:39)
ACATGACAGCAGATCCGCGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCTCTTTGCCCTCAT
CCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGCTGCCCCCAGGCTGGGTGTCC
CTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTTGCCCTGAGACAGCAGAATGTGGAAAGAC
35   TCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAGCTCTCCTCAATACGGAAAATACCTGACCCTAGAGAA
TGTGGCTGATCTGGTGAGGCCATCCCCACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGA
GCCCAGAAGTGCCATTCTGTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGC
TGCTGCTCCCTGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCC
ACATCCCTACCAGCTTCCACAGGCCTTGGCCCCCCATGTGGACTTTGTGGGGGGACTGCACCGTTTTCCC
40   CCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCTGCATCTGGGGGTAA
CCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGGGCTCTGGCACCAGCAATAACAG
CCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGACTCAGACCTGGCTCAGTTCATGCGCCTCTTC
GGTGGCAACTTTGCACATCAGGCATCAGTAGCCCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGA
TTGAGGCCAGTCTAGATGTGCAGTACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAG
45   CCCTGGCCGGCATGAGGGACAGGAGCCCTTCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTG
CCACATGTGCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGACAGTGGGGCCGG
GTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAGCCCCTATGTCACCACA
GTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAATTGTTGACTATATCAGTGGTGGTG
50   GCTTCAGCAATGTGTTCCCACGGCCTTCATACCAGGAGGAAGCTGTAACGAAGTTCCTGAGCTCTAGCCC
CCACCTGCCACCATCCAGTTACTTCAATGCCAGTGGCCGTGCCTACCAGATGTGGCTGCACTTTCTGAT
GGCTACTGGGTGGTCAGCAACAGAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGT
TTGGGGGGATCCTATCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCCTCTTGGCTTTCTCAA
CCCAAGGCTCTACCAGCAGCATGGGGCAGGACTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTCTG
55   GATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCCTGGCTGGGATCCTGTAACAGGCTGGGGAACAC
CCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATCAGGAGAGATGGCTTGTC
CCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTGGAAGCCCTGCTGAACCCTCAACTAT
TGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAATGCGGTGAGCTTGACTTGACTCCCAACCCTACC
ATGCTCCATCATACTCAGGTCTCCCTACTCCTGCCTTAGATTCCTCAATAAGATGCTGTAACTAGCATTT
60   TTTGAATGCCTCTCCCTCCGCATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAG
ACTCTGTGCACTATTTCACTTGATATTCATTCCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTT

```
      ACTCTTTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATGGCC
      TTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTGACTACTCTTGT
      CTTCCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGCTCCATTTGTAGATTTTTGC
      TCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACATCTACAACCATTACCATCTCACTAAAT
  5   AAGACTTTCTATCCAATAATGATTGATACCTCAAATGTAAGATGCGTGATACTCAACATTTCATCGTCCA
      CCTTCCCAACCCCAAACAATTCCATCTCGTTTCTTCTTGGTAAATGATGCTATGCTTTTTCCAACCAAGC
      CAGAAACCTGTGTCATCTTTTCACCCCACCTTCAATCAACAAGTCCTCAATCAACAAGTCCTACTGACTG
      CACATCTTAAATATATCTTTATCAGTCCACAAGTCCTTCCAATTATATTTCCCAAGTATATCTAGAACTT
      ATCCACTTATATCCCCACTGCTACTACCTTAGTTTAGGGCTATATTCTCTTGAAAAAAAGTGTCCTTACT
 10   TCCTGCCAATCCCCAAGTCATCTTCCAGAGTAAAATGCAAATCCCATCAGGCCACTTGGATGAAAACCCT
      TCAAGGATTACTGGATAGAATTCAGGCTTTCCCCTCCAGCCCCCAATCATAGCTCACAAACCTTCCTTGC
      TATTTGTTCTTAAGTAAAAAATCATTTTTCCTCCTCCCTCCCCAAACCCCAAGGAACTCTCACTCTTGCT
      CAAGCTGTTCCGTCCCCTTACCACCCCTGATACAACTGCCAGGTTAATTTCCAGAATTCTTGCAAGACTC
      AGTTCAGAAGTCACCTTCTTTCGTGAATGTTTTGATTCCCTGAGGCTACTTTATTTTGGTATGGCTGAAA
 15   AATCCTAGATTTTCTAAACAAAACCTGTTTGAATCTTGGTTCTGATATGGACTAGGAGAGAGACTGGGTC
      AAGTAAGCTTATCTCCCTGAGGCTGTTTCCTCGTCTGTTAAGTGTGAATATCAATACCTGCCTTTCATAA
      TCACCAGGAATAAAGTGGAATAATGTTGATAACAGTGCTTGGCACCTGGAAGTAGGTGGCAGATGTTAA
      CGCCCTTCCTCCCTTGCACTGCGCCCCTGTGCCTACCTCTAGCATTGTAACGACCACATAGTATTGAAA
      TGGCCAGTTTACTTGTCTGCCTTCCTTTCCAAGACCGTTGGTGCCTAGAGGACTAGAATCGTGTCCTATT
 20   TAACTTTGTGTTCCCAGGTCCTAGCTCAGGAGTTGGCAAATAAGAATTAAATGTCTGCTACACCGAAACA
      AA
```

>gi|5729770|ref|NP_000382.3| ceroid-lipofuscinosis, neuronal 2, late
infantile (Jansky-Bielschowsky disease) [Homo sapiens] (SEQ ID NO:40)
```
 25   MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLSELVQAVSD
      PSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWLSIRQAELLLPGAEFHH
      YVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRPEPQVTGTVGLHLGVTPSVIRKRYN
      LTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFGGNFAHQASVARVVGQQGRGRAGIEASLDVQYL
 30   MSAGANISTWVYSSPGRHEGQEPFLQWLMLLSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAAR
      GLTLLFASGDSGAGCWSVSGRHQFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPS
      YQEEAVTKFLSSSPHLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINE
      HRILSGRPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPNFPALLKTL
      LNP
```

 35

P2X5b

>gi|3387943|gb|AF070573.1|AF070573 Homo sapiens clone 24793 ionotropic ATP
receptor P2X5b mRNA, complete cds (SEQ ID NO:41)
```
 40   GTCCGCAAGCCCGGCTGAGAGCGCGCC<u>ATG</u>GGGCAGGCGGGCTGCAAGGGGCTCTGCCTGTCGCTGTTCG
      ACTACAAGACCGAGAAGTATGTCATCGCCAAGAACAAGAAGGTGGGCCTGCTGTACCGGCTGCTGCAGGC
      CTCCATCCTGGCGTACCTGGTCGTATGGGTGTTCCTGATAAAGAAGGGTTACCAAGACGTCGACACCTCC
      CTGCAGAGTGCTGTCATCACCAAAGTCAAGGGCGTGGCCTTCACCAACACCTCGGATCTTGGGCAGCGGA
      TCTGGGATGTCGCCGACTACGTCATTCCAGCCCAGAATGAAGGCATTCCTGATGGCGCGTGCTCCAAGGA
 45   CAGCGACTGCCACGCTGGGGAAGCGGTTACAGCTGGAAACGGAGTGAAGACCGGCCGCTGCCTGCGGAGA
      GAGAACTTGGCCAGGGGCACCTGTGAGATCTTTGCCTGGTGCCCGTTGGAGACAAGCTCCAGGCCGGAGG
      AGCCATTCCTGAAGGAGGCCGAAGACTTCACCATTTTCATAAAGAACCACATCCGTTTCCCCAAATTCAA
      CTTCTCCAACAATGTGATGGACGTCAAGGACAGATCTTTCCTGAAATCATGCCACTTTGGCCCCAAGAAC
      CACTACTGCCCCATCTTCCGACTGGGCTCCGTGATCCGCTGGGCCGGGAGCGACTTCCAGGATATAGCCC
 50   TGGAGGGTGGCGTGATAGGAATTAATATTGAATGGAACTGTGATCTTGATAAAGCTGCCTCTGAGTGCCA
      CCCTCACTATTCTTTTAGCCGTCTGGACAATAAACTTTCAAAGTCTGTCTCCTCCGGGTACAACTTCAGA
      TTTGCCAGATATTACCGAGACGCAGCCGGGGTGGAGTTCCGCACCCTGATGAAAGCCTACGGGATCCGCT
      TTGACGTGATGGTGAACGGCAAGGGTGCTTTCTTCTGCGACCTGGTACTCATCTACCTCATCAAAAAGAG
      AGAGTTTTACCGTGACAAGAAGTACGAGGAAGTGAGGGGCCTAGAAGACAGTTCCCAGGAGGCCGAGGAC
 55   GAGGCATCGGGGCTGGGGCTATCTGAGCAGCTCACATCTGGGCCAGGGCTGCTGGGGATGCCGGAGCAGC
      AGGAGCTGCAGGAGCCACCCGAGGCGAAGCGTGGAAGCAGCAGTCAGAAGGGGAACGGATCTGTGTGCCC
      ACAGCTCCTGGAGCCCCACAGGAGCACG<u>TGA</u>ATTGCCTCTGCTTACGTTCAGGCCCTGTCCTAAACCCAG
      CCGTCTAGCACCCAGTGATCCCATGCCTTTGGGAATCCCAGGATGCTGCCCAACGGGAAATTTGTACATT
      GGGTGCTATCAATGCCACATCACAGGGACCAGCCATCACAGAGCAAAGTGACCTCCACGTCTGATGCTGG
 60   GGTCATCAGGACGGACCCATCATGGCTGTCTTTTGCCCCACCCCTGCCGTCAGTTCTTCCTTTCTCCG
      TGGCTGGCTTCCCGCACTAGGGAACGGGTTGTAAATGGGGAACATGACTTCCTTCCGGAGTCCTTGAGCA
```

```
CCTCAGCTAAGGACCGCAGTGCCCTGTAGAGTTCCTAGATTACCTCACTGGGAATAGCATTGTGCGTGTC
CGGAAAAGGGCTCCATTTGGTTCCAGCCCACTCCCCTCTGCAAGTGCCACAGCTTCCCTCAGAGCATACT
CTCCAGTGGATCCAAGTACTCTCTCTCCTAAAGACACCACCTTCCTGCCAGCTGTTTGCCCTTAGGCCAG
TACACAGAATTAAAGTGGGGGAGATGGCAGACGCTTTCTGGGACCTGCCCAAGATATGTATTCTCTGACA
CTCTTATTTGGTCATAAAACAATAAATGGTGTCAATTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA
```

>gi|3387944|gb|AAC28645.1| ionotropic ATP receptor P2X5b [Homo sapiens] (SEQ ID NO:42)

```
MGQAGCKGLCLSLFDYKTEKYVIAKNKKVGLLYRLLQASILAYLVVWVFLIKKGYQDVDTSLQSAVITKV
KGVAFTNTSDLGQRIWDVADYVIPAQNEGIPDGACSKDSDCHAGEAVTAGNGVKTGRCLRRENLARGTCE
IFAWCPLETSSRPEEPFLKEAEDFTIFIKNHIRFPKFNFSNNVMDVKDRSFLKSCHFGPKNHYCPIFRLG
SVIRWAGSDFQDIALEGGVIGINIEWNCDLDKAASECHPHYSFSRLDNKLSKSVSSGYNFRFARYYRDAA
GVEFRTLMKAYGIRFDVMVNGKGAFFCDLVLIYLIKKREFYRDKKYEEVRGLEDSSQEAEDEASGLGLSE
QLTSGPGLLGMPEQQELQEPPEAKRGSSSQKGNGSVCPQLLEPHRST
```

6-PFKL

>gi|35397|emb|X16914.1|HSPFK04 Human PFKL gene for liver-type 6-phosphofructokinase (EC 2.7.1.11) exon 4 (SEQ ID NO:43)

```
CCAGTCCTGGGTCCCTCTGGTGATCCCAGGGCTGTCTGCCGCCTGCCATCTCTCCTGAAGTTTCTGGTCT
CCTCTGTGCAGGGCGCGACTATCATTGGCACGGCTCGCTCGAAGGCCTTTACCACCAGGGAGGGGCGCCG
GGCAGCGGCTAACAACCTGGTCCAGCACGGCATCACCAACCTGTGCGTCATCGGCGGGGATGGCAGCCTT
ACAGGTGCCAACATCTTCCGCAGCGAGTGGGGCAGCCTGCTGGAGGAGCTGGTGGCGGAAGGTGGGTCTG
TGCCCGGCGCACTGTAGGC
```

DUSP1

>gi|7108342|ref|NM_004417.2| Homo sapiens dual specificity phosphatase 1 (DUSP1), mRNA (SEQ ID NO:44)

```
TCGCTGCGAAGGACATTTGGGCTGTGTGTGCGACGCGGGTCGGAGGGGCAGTCGGGGGAACCGCGAAGAA
GCCGAGGAGCCCGGAGCCCCGCGTGACGCTCCTCTCTCAGTCCAAAAGCGGCTTTTGGTTCGGCGCAGAG
AGACCCGGGGGTCTAGCTTTTCCTCGAAAAGCGCCGCCCTGCCCTTGGCCCCGAGAACAGACAAAGAGCA
CCGCAGGGCCGATCACGCTGGGGGCGCTGAGGCCGGCCATGGTCATGGAAGTGGGCACCCTGGACGCTGG
AGGCCTGCGGGCGCTGCTGGGGGAGCGAGCGGCGCAATGCCTGCTGCTGGACTGCCGCTCCTTCTTCGCT
TTCAACGCCGGCCACATCGCCGGCTCTGTCAACGTGCGCTTCAGCACCATCGTGCGGCGCCGGGCCAAGG
GCGCCATGGGCCTGGAGCACATCGTGCCCAACGCCGAGCTCCGCGGCCGCCTGCTGGCCGGCGCCTACCA
CGCCGTGGTGTTGCTGGACGAGCGCAGCGCCGCCCTGGACGGCGCCAAGCGCGACGGCACCCTGGCCCTG
GCGGCCGGCGCGCTCTGCCGCGAGGCGCGCGCCGCGCAAGTCTTCTTCCTCAAAGGAGGATACGAAGCGT
TTTCGGCTTCCTGCCCGGAGCTGTGCAGCAAACAGTCGACCCCCATGGGGCTCAGCCTTCCCCTGAGTAC
TAGCGTCCCTGACAGCGCGGAATCTGGGTGCAGTTCCTGCAGTACCCCACTCTACGATCAGGGTGGCCCG
GTGGAAATCCTGCCCTTTCTGTACCTGGGCAGTGCGTATCACGCTTCCCGCAAGGACATGCTGGATGCCT
TGGGCATAACTGCCTTGATCAACGTCTCAGCCAATTGTCCCAACCATTTTGAGGGTCACTACCAGTACAA
GAGCATCCCTGTGGAGGACAACCACAAGGCAGACATCAGCTCCTGGTTCAACGAGGCCATTGACTTCATA
GACTCCATCAAGAATGCTGGAGGAAGGGTGTTTGTCCACTGCCAGGCAGGCATTTCCCGGTCAGCCACCA
TCTGCCTTGCTTACCTTATGAGGACTAATCGAGTCAAGCTGGACGAGGCCTTTGAGTTTGTGAAGCAGAG
GCGAAGCATCATCTCTCCCAACTTCAGCTTCATGGGCCAGCTGCTGCAGTTTGAGTCCCAGGTGCTGGCT
CCGCACTGTTCGGCAGAGGCTGGGAGCCCCGCCATGGCTGTGCTCGACCGAGGCACCTCCACCACCACCG
TGTTCAACTTCCCCGTCTCCATCCCTGTCCACTCCACGAACAGTGCGCTGAGCTACCTTCAGAGCCCCAT
TACGACCTCTCCCAGCTGCTGAAAGGCCACGGGAGGTGAGGCTCTTCACATCCCATTGGGACTCCATGCT
CCTTGAGAGGAGAAATGCAATAACTCTGGGAGGGGCTCGAGAGGGCTGGTCCTTATTTATTTAACTTCAC
CCGAGTTCCTCTGGGTTTCTAAGCAGTTATGGTGATGACTTAGCGTCAAGACATTTGCTGAACTCAGCAC
ATTCGGGACCAATATATAGTGGGTACATCAAGTCCATCTGACAAAATGGGGCAGAAGAGAAAGGACTCAG
TGTGTGATCCGGTTTCTTTTTGCTCGCCCCTGTTTTTTGTAGAATCTCTTCATGCTTGACATACCTACCA
GTATTATTCCCGACGACACATATACATATGAGAATATACCTTATTTATTTTTGTGTAGGTGTCTGCCTTC
ACAAATGTCATTGTCTACTCCTAGAAGAACCAAATACCTCAATTTTTGTTTTTGAGTACTGTACTATCCT
GTAAATATATCTTAAGCAGGTTTGTTTTCAGCACTGATGGAAAATACCAGTGTTGGGTTTTTTTTTAGTT
GCCAACAGTTGTATGTTTGCTGATTATTTATGACCTGAAATAATATATTTCTTCTTCTAAGAAGACATTT
TGTTACATAAGGATGACTTTTTTATACAATGGAATAAATTATGGCATTTCTATTG
```

>gi|4758204|ref|NP_004408.1| dual specificity phosphatase 1;
serine/threonine specific protein phosphatase [Homo sapiens] (SEQ ID NO:45)
MVMEVGTLDAGGLRALLGERAAQCLLLDCRSFFAFNAGHIAGSVNVRFSTIVRRRAKGAMGLEHIVPNAE

5 LRGRLLAGAYHAVVLLDERSAALDGAKRDGTLALAAGALCREARAAQVFFLKGGYEAFSASCPELCSKQS
TPMGLSLPLSTSVPDSAESGCSSCSTPLYDQGGPVEILPFLYLGSAYHASRKDMLDALGITALINVSANC
PNHFEGHYQYKSIPVEDNHKADISSWFNEAIDFIDSIKNAGGRVFVHCQAGISRSATICLAYLMRTNRVK
LDEAFEFVKQRRSIISPNFSFMGQLLQFESQVLAPHCSAEAGSPAMAVLDRGTSTTTVFNFPVSIPVHST
NSALSYLQSPITTSPSC

10

KIAA0251

>gi|2055294|dbj|D87438.1| Human mRNA for KIAA0251 gene, partial cds (SEQ ID
15 NO:46)
CGGGGGACGTCAGCGCTGCCAGCGTGGAAGGAGCTGCGGGGCGCGGGAGGAGGAAGTAGAGCCCGGGACC
GCCAGGCCACCACCGGCCGCCTCAGCCATGGACGCGTCCCTGGAGAAGATAGCAGACCCCACGTTAGCTG
AAATGGGAAAAAACTTGAAGGAGGCAGTGAAGATGCTGGAGGACAGTCAGAGAAGAACAGAAGAGGAAAA
TGGAAAGAAGCTCATATCCGGAGATATTCCAGGCCCACTCCAGGGCAGTGGGCAAGATATGGTGAGCATC
20 CTCCAGTTAGTTCAGAATCTCATGCATGGAGATGAAGATGAGGAGCCCCAGAGCCCCAGAATCCAAAATA
TTGGAGAACAAGGTCATATGGCTTTGTTGGGACATAGTCTGGGAGCTTATATTTCAACTCTGGACAAAGA
GAAGCTGAGAAAACTTACAACTAGGATACTTTCAGATACCACCTTATGGCTATGCAGAATTTTCAGATAT
GAAAATGGGTGTGCTTATTTCCACGAAGAGGAAAGAGAAGGACTTGCAAAGATATGTAGGCTTGCCATTC
ATTCTCGATATGAAGACTTCGTAGTGGATGGCTTCAATGTGTTATATAACAAGAAGCCTGTCATATATCT
25 TAGTGCTGCTGCTAGACCTGGCCTGGGCCAATACCTTTGTAATCAGCTCGGCTTGCCCTTCCCTGCTTG
TGCCGTGTACCCTGTAACACTGTGTTTGGATCCCAGCATCAGATGGATGTTGCCTTCCTGGAGAAACTGA
TTAAAGATGATATAGAGCGAGGAAGACTGCCCCTGTTGCTTGTCGCAAATGCAGGAACGGCAGCAGTAGG
ACACACAGACAAGATTGGGAGATTGAAAGAACTCTGTGAGCAGTATGGCATATGGCTTCATGTGGAGGGT
GTGAATCTGGCAACATTGGCTCTGGGTTATGTCTCCTCATCAGTGCTGGCTGCAGCCAAATGTGATAGCA
30 TGACGATGACTCCTGGCCCGTGGCTGGGTTTGCCAGCTGTTCCTGCGGTGACACTGTATAAACACGATGA
CCCTGCCTTGACTTTAGTTGCTGGTCTTACATCAAATAAGCCCACAGACAAACTCCGTGCCCTGCCTCTG
TGGTTATCTTTACAATACTTGGGACTTGATGGGTTTGTGGAGAGGATCAAGCATGCCTGTCAACTGAGTC
AACGGTTGCAGGAAAGTTTGAAGAAAGTGAATTACATCAAAATCTTGGTGGAAGATGAGCTCAGCTCCCC
AGTGGTGGTGTTCAGATTTTTCCAGGAATTACCAGGCTCAGATCCGGTGTTTAAAGCCGTCCCAGTGCCC
35 AACATGACACCTTCAGGAGTCGGCCGGGAGAGGCACTCGTGTGACGCGCTGAATCGCTGGCTGGGAGAAC
AGCTGAAGCAGCTGGTGCCTGCAAGCGGCCTCACAGTCATGGATCTGGAAGCTGAGGGCACGTGTTTGCG
GTTCAGCCCTTTGATGACCGCAGCAGTTTTAGGAACTCGGGGAGAGGATGTGGATCAGCTCGTAGCCTGC
ATAGAAAGCAAACTGCCAGTGCTGTGCTGTACGCTCCAGTTGCGTGAAGAGTTCAAGCAGGAAGTGGAAG
CAACAGCAGGTCTCCTATATGTTGATGACCCTAACTGGTCTGGAATAGGGGTTGTCAGGTATGAACATGC
40 TAATGATGATAAGAGCAGTTTGAAATCAGATCCCGAAGGGGAAAACATCCATGCTGGACTCCTGAAGAAG
TTAAATGAACTGGAATCTGACCTAACCTTTAAAATAGGCCCTGAGTATAAGAGCATGAAGAGCTGCCTTT
ATGTCGGCATGGCGAGCGACAACGTCGATGCTGCTGAGCTCGTGGAGACCATTGCGGCCACAGCCCGGGA
GATAGAGGAGAACTCGAGGCTTCTGGAAAACATGACAGAAGTGGTTCGGAAAGGCATTCAGGAAGCTCAA
GTGGAGCTGCAGAAGGCAAGTGAAGAACGGCTTCTGGAAGAGGGGGTGTTGCGGCAGATCCCTGTAGTGG
45 GCTCCGTGCTGAATTGGTTTTCTCCGGTCCAGGCTTTACAGAAGGGAAGAACTTTTAACTTGACAGCAGG
CTCTCTGGAGTCCACAGAACCCATATATGTCTACAAAGCACAAGGTGCAGGAGTCACGCTGCCTCCAACG
CCCTCGGGCAGTCGCACCAAGCAGAGGCTTCCAGGCCAGAAGCCTTTTAAAAGGTCCCTGCGAGGTTCAG
ATGCTTTGAGTGAGACCAGCTCAGTCAGTCACATTGAAGACTTAGAAAAGGTGGAGCGCCTATCCAGTGG
GCCGGAGCAGATCACCCTCGAGGCCAGCAGCACTGAGGGACACCCAGGGGCTCCCAGCCCTCAGCACACC
50 GACCAGACCGAGGCCTTCCAGAAAGGGGTCCCACACCCAGAAGATGACCACTCACAGGTAGAAGGACCGG
AGAGCTTAAGATGAGACTCATTGTGTGGTTTGAGACTGTACTGAGTATTGTTTCAGGGAAGATGAAGTTC
TATTGGAAATGTGAACTGTGCCACATACTAATATAAATTACTGTTGTTTGTGCTTCACTGGGATTTTGGC
ACAAATATGTGCCTGAAAGGTAGGCTTTCTAGGAGGGGAGTCAGCTTGTCTAACTTCATGTACATGTAGA
ACCACGTTTGCTGTCCTACTACGACTTTTCCCTAAGTTACCATAAACACATTTTATTCACAAAAAACACT
55 TCGAATTTCAAGTGTCTACCAGTAGCACCCTTGCTCTTTCTAAACATAAGCCTAAGTATATGAGGTTGCC
CGTGGCAACTTTTTGGTAAAACAGCTTTTCATTAGCACTCTCCAGGTTCTCTGCAACACTTCACAGAGGC
GAGACTGGCTGTATCCTTTGCTGTCGGTCTTTAGTACGATCAAGTTGCAATATACAGTGGGACTGCTAGA
CTTGAAGGAGAGCAGTGATTGTGGGATTGTAAATAAGAGCATCAGAAGCCCTCCCCAGCTACTGCTCTTC
GTGGAGACTTAGTAAGGACTGTGTCTACTTGAGCTGTGGCAAGGCTGCTGTCTGGGACTGTCCTCTGCCA
60 CAAGGCCATTTCTCCCATTATATACCGTTTGTAAAGAGAAACTGTAAAGTCTCCTCCTGACCATATATTT
TTAAATACTGGCAAAGCTTTTAAAATTGGCACACAAGTACAGACTGTGCTCATTTCTGTTTAGTATCTGA

```
        AAACCTGATAGATGCTACCCTTAAGAGCTTGCTCTTCCGTGTGCTACGTAGCACCCACCTGGTTAAAATC
        TGAAAACAAGTACCCCTTTGACCTGTCTCCCACTGAAGCTTCTACTGCCCTGGCAGCTCGCCTGGGCCCA
        ACTCAGAAACAGGAGCCAGCAGAGCACTCTCTCACGCTGATCCAGCCGGGCACCCTGCTTAAGTCAGTAG
        AAGCTCGCTGGCACTGCCCGTTCCTACTTTTCCGAAGTACTGCGTCACTTTGTCGTAAGTAATGGCCCCT
5       GTGCCTTCTTAATCCAGCAGTCAAGCTTTTGGGAGACCTGAAAATGGGAAAATTCACACTGGGTTTCTGG
        ACTGTAGTATTGGAAGCCTTAGTTATAGTATATTAAGCCTATAATTATACTCTGATTTGATGGGATTTTT
        GACATTTACACTTGTCAAAATGCAGGGGGTTTTTTTTGGTGCAGATGATTAAACAGTCTTCCCTATTTGG
        TGCAATGAAGTATAGCAGATAAAATGGGGGAGGGGTAAATTATCACCTTCAAGAAAATTACATGTTTTTA
        TATATATTTGGAATTGTTAAATTGGTTTTGCTGAAACATTTCACCCTTGAGATATTATTTGAATGTTGGT
10      TTCAATAAAGGTTCTTGAAATTGTT
```

>gi|2055295|dbj|BAA19780.1| Similar to a C. elegans protein in cosmid
C14H10 [Homo sapiens] (SEQ ID NO:47)

```
        GGRQRCQRGRSCGAREEEVEPGTARPPPAASAMDASLEKIADPTLAEMGKNLKEAVKMLEDSQRRTEEEN
15      GKKLISGDIPGPLQGSGQDMVSILQLVQNLMHGDEDEEPQSPRIQNIGEQGHMALLGHSLGAYISTLDKE
        KLRKLTTRILSDTTLWLCRIFRYENGCAYFHEEEREGLAKICRLAIHSRYEDFVVDGFNVLYNKKPVIYL
        SAAARPGLGQYLCNQLGLPFPCLCRVPCNTVFGSQHQMDVAFLEKLIKDDIERGRLPLLLVANAGTAAVG
        HTDKIGRLKELCEQYGIWLHVEGVNLATLALGYVSSSVLAAAKCDSMTMTPGPWLGLPAVPAVTLYKHDD
        PALTLVAGLTSNKPTDKLRALPLWLSLQYLGLDGFVERIKHACQLSQRLQESLKKVNYIKILVEDELSSP
20      VVVFRFFQELPGSDPVFKAVPVPNMTPSGVGRERHSCDALNRWLGEQLKQLVPASGLTVMDLEAEGTCLR
        FSPLMTAAVLGTRGEDVDQLVACIESKLPVLCCTLQLREEFKQEVEATAGLLYVDDPNWSGIGVVRYEHA
        NDDKSSLKSDPEGENIHAGLLKKLNELESDLTFKIGPEYKSMKSCLYVGMASDNVDAAELVETIAATARE
        IEENSRLLENMTEVVRKGIQEAQVELQKASEERLLEEGVLRQIPVVGSVLNWFSPVQALQKGRTFNLTAG
        SLESTEPIYVYKAQGAGVTLPPTPSGSRTKQRLPGQKPFKRSLRGSDALSETSSVSHIEDLEKVERLSSG
25      PEQITLEASSTEGHPGAPSPQHTDQTEAFQKGVPHPEDDHSQVEGPESLR
```


GG2-1

>gi|3978237|gb|AF070671.1|AF070671 Homo sapiens TNF-induced protein GG2-1
mRNA, complete cds (SEQ ID NO:48)

```
        CGGGAACCCGTGAGCCACCGAGAGAGCAGAGAACTCGGCGCCGCCAAACAGCCCAGCTCGCGCTTCAGCG
        TCCCGGCGCCGTCGCGCCACTCCTCCGATGGCCACAGATGTCTTTAATTCCAAAAACCTGGCCGTTCAGG
        CACAAAAGAAGATCTTGGGTAAAATGGTGTCCAAATCCATCGCCACCACCTTAATAGACGACACAAGTAG
        TGAGGTGCTGGATGAGCTCTACAGAGTGACCAGGGAGTACACCCAAAACAAGAAGGAGGCAGAGAAGATC
35      ATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGGAATAATCAGTTTAATCAAGATG
        AGCTAGCATTGATGGAGAAATTTAAGAAGAAAGTTCATCAGCTTGCTATGACCGTGGTCAGTTTCCATCA
        GGTGGATTATACCTTTGACCGGAATGTGTTATCCAGGCTGTTAAATGAATGCAGAGAGATGCTGCACCAA
        ATCATTCAGCGCCACCTCACTGCCAAGTCACATGGACGGGTTAATAATGTCTTTGATCATTTTTCAGATT
40      GTGAATTTTTGGCTGCCTTGTATAATCCTTTTGGGAATTTTAAACCCCACTTACAAAAACTATGTGATGG
        TATCAACAAAATGTTGGATGAAGAGAACATATGAGCACATGAGTTAAGATTGTGACTGATCATGATTTAT
        TTGAAGATGGAGCACTGCTGATTTATGAAGGAAAAAAGAAGAATTTTCTAAAGATTACACATATTTCAGA
        AAGACTTTACCCAATTCAGTTGTCAGACATAATGATTTATTTGAAGGCTTGTTTTATTTGAAGAAAGCA
        TATTGCCAAAAATTCTGGTTAAAAGCTTCCTAATGGGTAACAGACCATGGGAGAGATATGTGGTTGGGTA
45      ATGCGAATGTAGTTATACAAAGAAAAATACAGATGTCTCCAGACCTGAGGACTTTTTAATAGGGCAGTTG
        TTGTGTTGGTGGCACATTGGATATTTCTAACATGTACAAAGCTATGTATTTGATTTACTTTCATTTCTT
        GCTATGTATATGTACTTTTCTTAAAATGCCAAGAACTTTCTCTTGCTATCATTGCTCCTTTTGAAACAAT
        TCAATTTTCATGTCTACAGCTGACTGTTTTGTTAAGATTGAGTCATCGACATTCAGGATTTAAGTCTGAG
        GTAGTCAACCCTCAGGAAAAAAAAATGGCTTATCTGAAATCAGTACTGTGGAAATGAACTATATTAGCT
50      ATTATGAATAATGTCCAGTATAAGAATATGCTTCTGGAATTGAGTTCTCCTTTTAAGTACCAATGATACT
        TAAATTTCTCAGAAATGTAATGGTGTGTCATTGCCTTGAAATGCTTGCTTAGGGCTTCTTTTATGTTATC
        TTAAAAAGTGCTGGTGAATTTTCCATTTTTTACATCCATTTCACATGTAAGAGACAAAAAAGTCTAGATT
        GGTCTTGATATTGAGATAATAAAAAGTAAGTAGCATTAAGAAAGGTAACAATCTTCATTCTACAGATGAA
        CTCATTGAAACAATTTAGGGGAATGAGGGGCAAAAGGGGAGAAATACTGCTAAAGAACATGAGCATAAAA
55      ATGCGTGCGTTTCAGTGTTTAAGAAGGCTTGATAAAGAATGTCACTTTTTTATTTAACTGATAAGATTTT
        TGTTATTTTTTACTTTGATAAGTAAACCAAAGAATATTTGTATTTCAAGCAGTTTGTGTGGTGTTTCTAT
        ATAATTTTCTGTGTATAAATAATAAAGTAGGCATTTGTTTATTTTGTAAAAAGAAATGAAAATCTGCTG
        GCCAGCTATGTCCTCTAGGAAATGACAGACCCAACCACCAGCAATAAACATTTCCATTGCCAAAAAAAAA
        AA
```

60
```
                                        84
```

>gi|3978238|gb|AAC83229.1| TNF-induced protein GG2-1 [Homo sapiens] (SEQ ID
NO:49)
MATDVFNSKNLAVQAQKKILGKMVSKSIATTLIDDTSSEVLDELYRVTREYTQNKKEAEKIIKNLIKTVI
KLAILYRNNQFNQDELALMEKFKKKVHQLAMTVVSFHQVDYTFDRNVLSRLLNECREMLHQIIQRHLTAK
SHGRVNNVFDHFSDCEFLAALYNPFGNFKPHLQKLCDGINKMLDEENI

Grb7

>gi|4885354|ref|NM_005310.1| Homo sapiens growth factor receptor-bound
protein 7 (GRB7), mRNA (SEQ ID NO:50)
CACAGGGCTCCCCCCCGCCTCTGACTTCTCTGTCCGAAGTCGGGACACCCTCCTACCACCTGTAGAGAAG
CGGGAGTGGATCTGAAATAAAATCCAGGAATCTGGGGGTTCCTAGACGGAGCCAGACTTCGGAACGGGTG
TCCTGCTACTCCTGCTGGGGCTCCTCCAGGACAAGGGCACACAACTGGTTCCGTTAAGCCCCTCTCTCGC
TCAGACGCCATGGAGCTGGATCTGTCTCCACCTCATCTTAGCAGCTCTCCGGAAGACCTTTGGCCAGCCC
CTGGGACCCCTCCTGGGACTCCCCGGCCCCCTGATACCCCTCTGCCTGAGGAGGTAAAGAGGTCCCAGCC
TCTCCTCATCCCAACCACCGGCAGGAAACTTCGAGAGGAGGAGAGGCGTGCCACCTCCCTCCCCTCTATC
CCCAACCCCTTCCCTGAGCTCTGCAGTCCTCCCTCACAGAGCCCAATTCTCGGGGGCCCCTCCAGTGCAA
GGGGGCTGCTCCCCCGCGATGCCAGCCGCCCCCATGTAGTAAAGGTGTACAGTGAGGATGGGGCCTGCAG
GTCTGTGGAGGTGGCAGCAGGTGCCACAGCTCGCCACGTGTGTGAAATGCTGGTGCAGCGAGCTCACGCC
TTGAGCGACGAGACCTGGGGGCTGGTGGAGTGCCACCCCCACCTAGCACTGGAGCGGGGTTTGGAGGACC
ACGAGTCCGTGGTGGAAGTGCAGGCTGCCTGGCCCGTGGGCGGAGATAGCCGCTTCGTCTTCCGGAAAAA
CTTCGCCAAGTACGAACTGTTCAAGAGCTCCCCACACTCCCTGTTCCCAGAAAAAATGGTCTCCAGCTGT
CTCGATGCACACACTGGTATATCCCATGAAGACCTCATCCAGAACTTCCTGAATGCTGGCAGCTTTCCTG
AGATCCAGGGCTTTCTGCAGCTGCGGGGTTCAGGACGGAAGCTTTGGAAACGCTTTTTCTGTTTCTTGCG
CCGATCTGGCCTCTATTACTCCACCAAGGGCACCTCTAAGGATCCGAGGCACCTGCAGTACGTGGCAGAT
GTGAACGAGTCCAACGTGTACGTGGTGACGCAGGGCCGCAAGCTCTACGGGATGCCCACTGACTTCGGTT
TCTGTGTCAAGCCCAACAAGCTTCGAAATGGACACAAGGGGCTTCGGATCTTCTGCAGTGAAGATGAGCA
GAGCCGCACCTGCTGGCTGGCTGCCTTCCGCCTCTTCAAGTACGGGGTGCAGCTGTACAAGAATTACCAG
CAGGCACAGTCTCGCCATCTGCATCCATCTTGTTTGGGCTCCCCACCCTTGAGAAGTGCCTCAGATAATA
CCCTGGTGGCCATGGACTTCTCTGGCCATGCTGGGCGTGTCATTGAGAACCCCGGGAGGCTCTGAGTGT
GGCCCTGGAGGAGGCCCAGGCCTGGAGGAAGAAGACAAACCACCGCCTCAGCCTGCCCATGCCAGCCTCC
GGCACGAGCCTCAGTGCAGCCATCCACCGCACCCAACTCTGGTTCCACGGGCGCATTTCCGTGAGGAGA
GCCAGCGGCTTATTGGACAGCAGGGCTTGGTAGACGGCCTGTTCCTGGTCCGGGAGAGTCAGCGGAACCC
CCAGGGCTTTGTCCTCTCTTTGTGCCACCTGCAGAAAGTGAAGCATTATCTCATCCTGCCGAGCGAGGAG
GAGGGTCGCCTGTACTTCAGCATGGATGATGGCCAGACCCGCTTCACTGACCTGCTGCAGCTCGTGGAGT
TCCACCAGCTGAACCGCGGCATCCTGCCGTGCTTGCTGCGCCATTGCTGCACGCGGGTGGCCCTCTGACC
AGGCCGTGGACTGGCTCATGCCTCAGCCCGCCTTCAGGCTGCCCGCCGCCCCTCCACCCATCCAGTGGAC
TCTGGGGCGCGGCCACAGGGACGGGATGAGGAGCGGGAGGGTTCCGCCACTCCAGTTTTCTCCTCTGCT
TCTTTGCCTCCCTCAGATAGAAAACAGCCCCCACTCCAGTCCACTCCTGACCCCTCTCCTCAAGGGAAGG
CCTTGGGTGGCCCCCTCTCCTTCTCCTAGCTCTGGAGGTGCTGCTCTAGGGCAGGGAATTATGGGAGAAG
TGGGGGCAGCCCAGGCGGTTTCACGCCCCACACTTTGTACAGACCGAGAGGCCAGTTGATCTGCTCTGTT
TTATACTAGTGACAATAAAGATTATTTTTTGATAC

>gi|4885355|ref|NP_005301.1| growth factor receptor-bound protein 7 [Homo
sapiens] (SEQ ID NO:51)
MELDLSPPHLSSSPEDLWPAPGTPPGTPRPPDTPLPEEVKRSQPLLIPTTGRKLREEERRATSLPSIPNP
FPELCSPPSQSPILGGPSSARGLLPRDASRPHVVKVYSEDGACRSVEVAAGATARHVCEMLVQRAHALSD
ETWGLVECHPHLALERGLEDHESVVEVQAAWPVGGDSRFVFRKNFAKYELFKSSPHSLFPEKMVSSCLDA
HTGISHEDLIQNFLNAGSFPEIQGFLQLRGSGRKLWKRFFCFLRRSGLYYSTKGTSKDPRHLQYVADVNE
SNVYVVTQGRKLYGMPTDFGFCVKPNKLRNGHKGLRIFCSEDEQSRTCWLAAFRLFKYGVQLYKNYQQAQ
SRHLHPSCLGSPPLRSASDNTLVAMDFSGHAGRVIENPREALSVALEEAQAWRKKTNHRLSLPMPASGTS
LSAAIHRTQLWFHGRISREESQRLIGQQGLVDGLFLVRESQRNPQGFVLSLCHLQKVKHYLILPSEEEGR
LYFSMDDGQTRFTDLLQLVEFHQLNRGILPCLLRHCCTRVAL

SH2-B

>gi|8163910|gb|AF227968.1|AF227968 Homo sapiens SH2-B beta signaling
protein (SH2B) mRNA, complete cds, alternatively spliced (SEQ ID NO:52)
GAGCCGCCGCCGCCGCCGGAGCTAACCTCGGGGACCGAGATGCAGCTGCTGCCGCCCACCCCTCGTCTTC

```
TGGCTGCCTCCCTCTTTGTGCCCCACAGGCTCCCCCTCTCCACCTCCTGGGGCCCATCATGAATGGTGCC
CCTTCCCCAGAGGACGGGGCCTCCCCCTCGTCTCCCCGCTGCCCCCACCCCCGCCCCCTAGTTGGCGGG
AGTTCTGTGAGTCCCACGCCCGGGCTGCGGCTCTGGACTTTGCCCGCCGTTTTCGCCTCTACCTGGCCTC
CCACCCCCAATATGCGGGGCCCGGGGCCGAGGCTGCCTTCTCCCGCCGTTTTGCTGAGCTCTTCCTGCAG
CACTTTGAAGCCGAGGTGGCCCGGGCCTCTGGCTCCCTGTCGCCACCCATCCTGGCTCCCCTGAGCCCTG
GTGCGGAGATTTCGCCACATGACCTGTCCCTTGAGAGCTGCAGGGTGGGTGGGCCCCTGGCTGTGCTGGG
CCCTTCTCGATCATCTGAGGACCTGGCCGGCCCCCTCCCTTCCTCAGTCTCTTCCTCCTCTACAACCTCC
TCCAAGCCGAAGCTCAAGAAGCGCTTTTCCCTGCGTTCAGTGGGTCGCTCTGTCCGAGGCTCAGTCCGTG
GCATCCTGCAGTGGCGGGGGACCGTTGACCCTCCCTCCTCCGCTGGGCCCCTGGAGACCTCGTCAGGCCC
CCCTGTCTTAGGTGGAAACAGCAACTCCAACTCCTCTGGCGGGGCTGGGACCGTTGGTAGGGGACTGGTC
AGTGATGGAACGTCCCCTGGGGAAAGATGGACTCACCGTTTTGAGAGGCTGAGACTCAGTCGGGGAGGGG
GCGCCTTGAAGGATGGAGCAGGGATGGTGCAGAGGGAAGAGCTGCTGAGTTTCATGGGGGCTGAGGAGGC
AGCCCCTGACCCAGCCGGAGTGGGCCGGGGAGGAGGGGTGGCTGGGCCTCCTTCAGGGGGAGGAGGGCAG
CCTCAGTGGCAGAAGTGTCGCCTGCTGCTTCGAAGTGAAGGAGAAGGAGGAGGAGGAAGTCGCCTGGAGT
TCTTTGTACCACCCAAGGCCTCTCGGCCCCGACTCAGCATCCCCTGCTCTTCTATCACAGACGTCCGGAC
AACCACAGCCCTGGAGATGCCTGACCGGGAGAACACGTTTGTGGTTAAGGTGGAAGGTCCATCCGAGTAT
ATCATGGAGACAGTGGATGCCCAGCATGTGAAGGCCTGGGTGTCTGACATCCAAGAATGCCTGAGCCCAG
GACCCTGCCCTGCCACCAGTCCCCGCCCCATGACCCTCCCTCTGGCCCCTGGGACCTCATTCCTTACAAG
GGAGAACACAGACAGCCTGGAGCTGTCCTGCCTGAATCACTCGGAGAGTTTACCCAGCCAGGACCTGCTG
CTTGGACCCAGCGAGAGCAATGACCGCCTGTCGCAGGGGGCATATGGGGGCCTCTCAGACCGCCCCTCGG
CATCCATCTCCCCCAGCTCTGCCTCCATTGCCGCCTCCATTTTGACTCGATGGAACTGCTTCCCCCAGA
GTTGCCCCCCGCATCCCCATTGAAGAGGGACCCCCAGCAGGGACAGTTCATCCCCTCTCAGCCCCCTAC
CCTCCCTTGGACACTCCGGAAACAGCCACAGGGTCCTTCCTGTTCCAGGGGGAGCCAGAGGGCGGTGAGG
GGGACCAGCCCCTCTCAGGGTATCCTTGGTTCCACGGGATGCTCTCTCGGCTCAAGGCTGCACAGTTGGT
GCTGACTGGCGGCACTGGCTCCCACGGTGTCTTCCTGGTGCGCCAGAGTGAGACAAGGCGGGGTGAATAC
GTCCTCACCTTCAACTTCCAGGGCAAGGCCAAGCACCTGCGTTTGTCGCTGAACGAGGAGGGTCAGTGCC
GGGTCCAGCACCTGTGGTTCCAGTCCATTTTCGATATGCTCGAGCACTTCCGGGTGCACCCCATCCCTTT
GGAGTCGGGAGGCTCCAGTGATGTTGTCCTTGTCAGCTATGTCCCATCCTCCCAGCGACAGCAGGGCCGG
GAGCAGGCTGGGAGCCATGCGGGGGTGTGCGAGGGAGATGGATGCCACCCCGATGCCTCCTGCACCCTCA
TGCCCTTCGGAGCGAGTGACTGTGTAACCGACCACCTCCCATGACCCACCCCAGCCCCTGAACCCCCTT
CATGGACAGATCCCCCACAGCCTGGGGCAGAAGAGGCGTCGAGGGCGCCAGAAGTGGCGGCAGCAGCAGC
CGCAGCAGCCAAAGAGAGGCAAGAGAAAGAGAAAGCGGGCGGTGGAGGGGTCCCGGAAGAGCTGGTCCCC
GTGGTTGAGCTGGTCCCCGTGGTTGAATTGGAAGAGGCCATAGCCCCAGGCTCAGAGGCCCAGGGCGCTG
GGTCTGGTGGGGACGCGGGGGTGCCCCCAATGGTGCAGCTGCAGCAGTCACCACTAGGGGGTGATGGAGA
GGAAGGGGGCCACCCCAGGGCCATTAACAACCAGTACTCCTTCGTGTGAGCCAACCCCACCCGCTCCACC
CTTTTTAAACCCCCCAGCCCTGCTCGTGAGATTGGGCTGGGTAGGGACAGAGGAGGCCGAAATCCCTCCC
CCATGCTTCCTGACCCTTGTTGGCCAAGGGCATCTTTGATGGTACAAGCAGAGGCTCGGGAGAGGCTCCC
GTCACACTACAGGTCCCCTCCCCAGGGCAGGGGATTTGGGCTCCATGAGCTCCTTGAGGGGCTCTTCT
GGTCAGCCCCACCCTGGGGGCCATTTCCCCATTAACTACCCCCAGCCCGAGGCAGGGTGAGGGGGAAGGG
CTGTCAGTTACATTAAGGTGGTTGTTGTTGTTGTTTAAACAAAATGGAGAAGCATAAATAAATAAAAAG
GTTTATCTCGGTTCAAAAAAAAA
```

>gi|8163911|gb|AAF73913.1|AF227968_1 SH2-B beta signaling protein [Homo sapiens] (SEQ ID NO:53)
MNGAPSPEDGASPSSPPLPPPPPPSWREFCESHARAAALDFARRFRLYLASHPQYAGPGAEAAFSRRFAE
LFLQHFEAEVARASGSLSPPILAPLSPGAEISPHDLSLESCRVGGPLAVLGPSRSSEDLAGPLPSSVSSS
STTSSKPKLKKRFSLRSVGRSVRGSVRGILQWRGTVDPPSSAGPLETSSGPPVLGGNSNSNSSGGAGTVG
RGLVSDGTSPGERWTHRFERLRLSRGGGALKDGAGMVQREELLSFMGAEEAAPDPAGVGRGGGVAGPPSG
GGGQPQWQKCRLLLRSEGEGGGGSRLEFFVPPKASRPRLSIPCSSITDVRTTTALEMPDRENTFVVKVEG
PSEYIMETVDAQHVKAWVSDIQECLSPGPCPATSPRPMTLPLAPGTSFLTRENTDSLELSCLNHSESLPS
QDLLLGPSESNDRLSQGAYGGLSDRPSASISPSSASIAASHFDSMELLPPELPPRIPIEEGPPAGTVHPL
SAPYPPLDTPETATGSFLFQGEPEGGEGDQPLSGYPWFHGMLSRLKAAQLVLTGGTGSHGVFLVRQSETR
RGEYVLTFNFQGKAKHLRLSLNEEGQCRVQHLWFQSIFDMLEHFRVHPIPLESGGSSDVVLVSYVPSSQR
QQGREQAGSHAGVCEGDGCHPDASCTLMPFGASDCVTDHLP

STAT1

>gi|6274551|ref|NM_007315.1| Homo sapiens signal transducer and activator of transcription 1, 91kD (STAT1), mRNA (SEQ ID NO:54)
ATTAAACCTCTCGCCGAGCCCCTCCGCAGACTCTGCGCCGGAAAGTTTCATTTGCTGTATGCCATCCTCG

AGAGCTGTCTAGGTTAACGTTCGCACTCTGTGTATATAACCTCGACAGTCTTGGCACCTAACGTGCTGTG
CGTAGCTGCTCCTTTGGTTGAATCCCCAGGCCCTTGTTGGGGCACAAGGTGGCAGG<u>ATG</u>TCTCAGTGGTA
CGAACTTCAGCAGCTTGACTCAAAATTCCTGGAGCAGGTTCACCAGCTTTATGATGAC<u>AG</u>TTTTCCCATG
GAAATCAGACAGTACCTGGCACAGTGGTTAGAAAAGCAAGACTGGGAGCACGCTGCCAATGATGTTTCAT
TTGCCACCATCCGTTTTCATGACCTCCTGTCACAGCTGGATGATCAATATAGTCGCTTTTCTTTGGAGAA
TAACTTCTTGCTACAGCATAACATAAGGAAAAGCAAGCGTAATCTTCAGGATAATTTTCAGGAAGACCCA
ATCCAGATGTCTATGATCATTTACAGCTGTCTGAAGGAAGAAAGGAAAATTCTGGAAAACGCCCAGAGAT
TTAATCAGGCTCAGTCGGGGAATATTCAGAGCACAGTGATGTTAGACAAACAGAAAGAGCTTGACAGTAA
AGTCAGAAATGTGAAGGACAAGGTTATGTGTATAGAGCATGAAATCAAGAGCCTGGAAGATTTACAAGAT
GAATATGACTTCAAATGCAAAACCTTGCAGAACAGAGAACACGAGACCAATGGTGTGGCAAAGAGTGATC
AGAAACAAGAACAGCTGTTACTCAAGAAGATGTATTTAATGCTTGACAATAAGAGAAAGGAAGTAGTTCA
CAAAATAATAGAGTTGCTGAATGTCACTGAACTTACCCAGAATGCCCTGATTAATGATGAACTAGTGGAG
TGGAAGCGGAGACAGCAGAGCGCCTGTATTGGGGGGCCGCCCAATGCTTGCTTGGATCAGCTGCAGAACT
GGTTCACTATAGTTGCGGAGAGTCTGCAGCAAGTTCGGCAGCAGCTTAAAAAGTTGGAGGAATTGGAACA
GAAATACACCTACGAACATGACCCTATCACAAAAAACAAACAAGTGTTATGGGACCGCACCTTCAGTCTT
TTCCAGCAGCTCATTCAGAGCTCGTTTGTGGTGGAAAGACAGCCCTGCATGCCAACGCACCCTCAGAGGC
CGCTGGTCTTGAAGACAGGGGTCCAGTTCACTGTGAAGTTGAGACTGTTGGTGAAATTGCAAGAGCTGAA
TTATAATTTGAAAGTCAAAGTCTTATTTGATAAAGATGTGAATGAGAGAAATACAGTAAAAGGATTTAGG
AAGTTCAACATTTTGGGCACGCACACAAAAGTGATGAACATGGAGGAGTCCACCAATGGCAGTCTGGCGG
CTGAATTTCGGCACCTGCAATTGAAAGAACAGAAAAATGCTGGCACCAGAACGAATGAGGGTCCTCTCAT
CGTTACTGAAGAGCTTCACTCCCTTAGTTTTGAAACCCAATTGTGCCAGCCTGGTTTGGTAATTGACCTC
GAGACGACCTCTCTGCCCGTTGTGGTGATCTCCAACGTCAGCCAGCTCCCGAGCGGTTGGGCCTCCATCC
TTTGGTACAACATGCTGGTGGCGGAACCCAGGAATCTGTCCTTCTTCCTGACTCCACCATGTGCACGATG
GGCTCAGCTTTCAGAAGTGCTGAGTTGGCAGTTTTCTTCTGTCACCAAAAGAGGTCTCAATGTGGACCAG
CTGAACATGTTGGGAGAGAAGCTTCTTGGTCCTAACGCCAGCCCCGATGGTCTCATTCCGTGGACGAGGT
TTTGTAAGGAAAATATAAATGATAAAAATTTTCCCTTCTGGCTTTGGATTGAAAGCATCCTAGAACTCAT
TAAAAAACACCTGCTCCCTCTCTGGAATGATGGGTGCATCATGGGCTTCATCAGCAAGGAGCGAGAGCGT
GCCCTGTTGAAGGACCAGCAGCCGGGGACCTTCCTGCTGCGGTTCAGTGAGAGCTCCCGGGAAGGGGCCA
TCACATTCACATGGGTGGAGCGGTCCCAGAACGGAGGCGAACCTGACTTCCATGCGGTTGAACCCTACAC
GAAGAAAGAACTTTCTGCTGTTACTTTCCCTGACATCATTCGCAATTACAAAGTCATGGCTGCTGAGAAT
ATTCCTGAGAATCCCCTGAAGTATCTGTATCCAAATATTGACAAAGACCATGCCTTTGGAAAGTATTACT
CCAGGCCAAAGGAAGCACCAGAGCCAATGGAACTTGATGGCCCTAAAGGAACTGGATATATCAAGACTGA
GTTGATTTCTGTGTCTGAAGTTCACCCTTCTAGACTTCAGACCACAGACAACCTGCTCCCCATGTCTCCT
GAGGAGTTTGACGAGGTGTCTCGGATAGTGGGCTCTGTAGAATTCGACAGTATGATGAACACAGTA<u>TAGA</u>
GCATGAATTTTTTTCATCTTCTCTGGCGACAGTTTTCCTTCTCATCTGTGATTCCCTCCTGCTACTCTGT
TCCTTCACATCCTGTGTTTCTAGGGAAATGAAAGAAAGGCCAGCAAATTCGCTGCAACCTGTTGATAGCA
AGTGAATTTTTTCTCTAACTCAGAAACATCAGTTACTCTGAAGGGCATCATGCATCTTACTGAAGGTAAAA
TTGAAAGGCATTCTCTGAAGAGTGGGTTTCACAAGTGAAAAACATCCAGATACACCCAAAGTATCAGGAC
GAGAATGAGGGTCCTTTGGGAAAGGAGAAGTTAAGCAACATCTAGCAAATGTTATGCATAAAGTCAGTGC
CCAACTGTTATAGGTTGTTGGATAAATCAGTGGTTATTTAGGGAACTGCTTGACGTAGGAACGGTAAATT
TCTGTGGGAGAATTCTTACATGTTTTCTTTGCTTTAAGTGTAACTGGCAGTTTTCCATTGGTTTACCTGT
GAAATAGTTCAAAGCCAAGTTTATATACAATTATATCAGTCCTCTTTCAAAGGTAGCCATCATGGATCTG
GTAGGGGGAAAATGTGTATTTTATTACATCTTTCACATTGGCTATTTAAAGACAAAGACAAATTCTGTTT
CTTGAGAAGAGAATATTAGCTTTACTGTTTGTTATGGCTTAATGACACTAGCTAATATCAATAGAAGGAT
GTACATTTCCAAATTCACAAGTTGTGTTTGATATCCAAAGCTGAATACATTCTGCTTTCATCTTGGTCAC
ATACAATTATTTTTACAGTTCTCCCAAGGGAGTTAGGCTATTCACAACCACTCATTCAAAAGTTGAAATT
AACCATAGATGTAGATAAACTCAGAAATTTAATTCATGTTTCTTAAATGGGCTACTTTGTCCTTTTTGTT
ATTAGGGTGGTATTTAGTCTATTAGCCACAAAATTGGGAAAGGAGTAGAAAAAGCAGTAACTGACAACTT
GAATAATACACCAGAGATAATATGAGAATCAGATCATTTCAAAACTCATTTCCTATGTAACTGCATTGAG
AACTGCATATGTTTCGCTGATATATGTGTTTTTCACATTTGCGAATGGTTCCATTCTCTCTCCTGTACTT
TTTCCAGACACTTTTTTGAGTGGATGATGTTTCGTGAAGTATACTGTATTTTTACCTTTTTTCCTTCCTTA
TCACTGACACAAAAGTAGATTAAGAGATGGGTTTGACAAGGTTCTTCCCTTTTACATACTGCTGTCTAT
GTGGCTGTATCTTGTTTTTCCACTACTGCTACCACAACTATATTATCATGCAAATGCTGTATTCTTCTTT
GGTGGAGATAAAGATTTCTTGAGTTTTGTTTTAAAATTAAAGCTAAGTATCTGTATTGCATTAAATATA
ATATCGACACAGTGCTTTCCGTGGCACTGCATACAATCTGAGGCCTCCTCTCTCAGTTTTTATATAGATG
GCGAGAACCTAAGTTTCAGTTGATTTTACAATTGAAATGACTAAAAAACAAAGAAGACAACATTAAAAAC
AATATTGTTTCTA

>gi|6274552|ref|NP_009330.1| signal transducer and activator of
transcription 1 isoform alpha; signal transducer and activator of
transcription-1; signal transducer and activator of transcription 1, 91kD;

transcription factor ISGF-3; transcription factor ISGF-3 components p91/p84
[Homo sapiens] (SEQ ID NO:55)
MSQWYELQQLDSKFLEQVHQLYDDSFPMEIRQYLAQWLEKQDWEHAANDVSFATIRFHDLLSQLDDQYSR
FSLENNFLLQHNIRKSKRNLQDNFQEDPIQMSMIIYSCLKEERKILENAQRFNQAQSGNIQSTVMLDKQK
ELDSKVRNVKDKVMCIEHEIKSLEDLQDEYDFKCKTLQNREHETNGVAKSDQKQEQLLLKKMYLMLDNKR
KEVVHKIIELLNVTELTQNALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKL
EELEQKYTYEHDPITKNQVLWDRTFSLFQQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVK
LQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTN
EGPLIVTEELHSLSFETQLCQPGLVIDLETTSLPVVVISNVSQLPSGWASILWYNMLVAEPRNLSFFLTP
PCARWAQLSEVLSWQFSSVTKRGLNVDQLNMLGEKLLGPNASPDGLIPWTRFCKENINDKNFPFWLWIES
ILELIKKHLLPLWNDGCIMGFISKERERALLKDQQPGTFLLRFSESSREGAITFTWVERSQNGGEPDFHA
VEPYTKKELSAVTFPDIIRNYKVMAAENIPENPLKYLYPNIDKDHAFGKYYSRPKEAPEPMELDGPKGTG
YIKTELISVSEVHPSRLQTTDNLLPMSPEEFDEVSRIVGSVEFDSMMNTV

>gi|21536299|ref|NM_007315.2| Homo sapiens signal transducer and activator
of transcription 1, 91kDa (STAT1), transcript variant alpha, mRNA (SEQ ID
NO:56)
AGCGGGGCGGGGCGCCAGCGCTGCCTTTTCTCCTGCCGGGTAGTTTCGCTTTCCTGCGCAGAGTCTGCGG
AGGGGCTCGGCTGCACCGGGGGGATCGCGCCTGGCAGACCCCAGACCGAGCAGAGGCGACCCAGCGCGCT
CGGGAGAGGCTGCACCGCCGCGCCCCCGCCTAGCCCTTCCGGATCCTGCGCGCAGAAAAGTTTCATTTGC
TGTATGCCATCCTCGAGAGCTGTCTAGGTTAACGTTCGCACTCTGTGTATATAACCTCGACAGTCTTGGC
ACCTAACGTGCTGTGCGTAGCTGCTCCTTTGGTTGAATCCCCAGGCCCTTGTTGGGGCACAAGGTGGCAG
GATGTCTCAGTGGTACGAACTTCAGCAGCTTGACTCAAAATTCCTGGAGCAGGTTCACCAGCTTTATGAT
GACAGTTTTCCCATGGAAATCAGACAGTACCTGGCACAGTGGTTAGAAAAGCAAGACTGGGAGCACGCTG
CCAATGATGTTTCATTTGCCACCATCCGTTTTCATGACCTCCTGTCACAGCTGGATGATCAATATAGTCG
CTTTTCTTTGGAGAATAACTTCTTGCTACAGCATAACATAAGGAAAAGCAAGCGTAATCTTCAGGATAAT
TTTCAGGAAGACCCAATCCAGATGTCTATGATCATTTACAGCTGTCTGAAGGAAGAAAGGAAAATTCTGG
AAAACGCCCAGAGATTTAATCAGGCTCAGTCGGGGAATATTCAGAGCACAGTGATGTTAGACAAACAGAA
AGAGCTTGACAGTAAAGTCAGAAATGTGAAGGACAAGGTTATGTGTATAGAGCATGAAATCAAGAGCCTG
GAAGATTTACAAGATGAATATGACTTCAAATGCAAAACCTTGCAGAACAGAGAACACGAGACCAATGGTG
TGGCAAAGAGTGATCAGAAACAAGAACAGCTGTTACTCAAGAAGATGTATTTAATGCTTGACAATAAGAG
AAAGGAAGTAGTTCACAAAATAATAGAGTTGCTGAATGTCACTGAACTTACCCAGAATGCCCTGATTAAT
GATGAACTAGTGGAGTGGAAGCGGAGACAGCAGAGCGCCTGTATTGGGGGGCCGCCCAATGCTTGCTTGG
ATCAGCTGCAGAACTGGTTCACTATAGTTGCGGAGAGTCTGCAGCAAGTTCGGCAGCAGCTTAAAAAGTT
GGAGGAATTGGAACAGAAATACACCTACGAACATGACCCTATCACAAAAAACAAACAAGTGTTATGGGAC
CGCACCTTCAGTCTTTTCCAGCAGCTCATTCAGAGCTCGTTTGTGGTGGAAAGACAGCCCTGCATGCCAA
CGCACCCTCAGAGGCCGCTGGTCTTGAAGACAGGGGTCCAGTTCACTGTGAAGTTGAGACTGTTGGTGAA
ATTGCAAGAGCTGAATTATAATTTGAAAGTCAAAGTCTTATTTGATAAAGATGTGAATGAGAGAAATACA
GTAAAAGGATTTAGGAAGTTCAACATTTTGGGCACGCACACAAAAGTGATGAACATGGAGGAGTCCACCA
ATGGCAGTCTGGCGGCTGAATTTCGGCACCTGCAATTGAAAGAACAGAAAAATGCTGGCACCAGAACGAA
TGAGGGTCCTCTCATCGTTACTGAAGAGCTTCACTCCCTTAGTTTTGAAACCCAATTGTGCCAGCCTGGT
TTGGTAATTGACCTCGAGACGACCTCTCTGCCCGTTGTGGTGATCTCCAACGTCAGCCAGCTCCCGAGCG
GTTGGGCCTCCATCCTTTGGTACAACATGCTGGTGGCGGAACCCAGGAATCTGTCCTTCTTCCTGACTCC
ACCATGTGCACGATGGGCTCAGCTTTCAGAAGTGCTGAGTTGGCAGTTTTCTTCTGTCACCAAAAGAGGT
CTCAATGTGGACCAGCTGAACATGTTGGGAGAGAAGCTTCTTGGTCCTAACGCCAGCCCCGATGGTCTCA
TTCCGTGGACGAGGTTTTGTAAGGAAAATATAAATGATAAAAATTTTCCCTTCTGGCTTTGGATTGAAAG
CATCCTAGAACTCATTAAAAAACACCTGCTCCCTCTCTGGAATGATGGGTGCATCATGGGCTTCATCAGC
AAGGAGCGAGAGCGTGCCCTGTTGAAGGACCAGCAGCCGGGGACCTTCCTGCTGCGGTTCAGTGAGAGCT
CCCGGGAAGGGGCCATCACATTCACATGGGTGGAGCGGTCCCAGAACGGAGGCGAACCTGACTTCCATGC
GGTTGAACCCTACACGAAGAAAGAACTTTCTGCTGTTACTTTCCCTGACATCATTCGCAATTACAAAGTC
ATGGCTGCTGAGAATATTCCTGAGAATCCCCTGAAGTATCTGTATCCAAATATTGACAAAGACCATGCCT
TTGGAAAGTATTACTCCAGGCCAAAGGAAGCACCAGAGCCAATGGAACTTGATGGCCCTAAAGGAACTGG
ATATATCAAGACTGAGTTGATTTCTGTGTCTGAAGTTCACCCTTCTAGACTTCAGACCACAGACAACCTG
CTCCCCATGTCTCCTGAGGAGTTTGACGAGGTGTCTCGGATAGTGGGCTCTGTAGAATTCGACAGTATGA
TGAACACAGTATAGAGCATGAATTTTTTTCATCTTCTCTGGCGACAGTTTTCCTTCTCATCTGTGATTCC
CTCCTGCTACTCTGTTCCTTCACATCCTGTGTTTCTAGGGAAATGAAAGAAAGGCCAGCAAATTCGCTGC
AACCTGTTGATAGCAAGTGAATTTTTCTCTAACTCAGAAACATCAGTTACTCTGAAGGGCATCATGCATC
TTACTGAAGGTAAAATTGAAAGGCATTCTCTGAAGAGTGGGTTTCACAAGTGAAAAACATCCAGATACAC
CCAAGTATCAGGACGAGAATGAGGGTCCTTTGGGAAAGGAGAAGTTAAGCAACATCTAGCAAATGTTAT
GCATAAAGTCAGTGCCCAACTGTTATAGGTTGTTGGATAAATCAGTGGTTATTTAGGGAACTGCTTGACG
TAGGAACGGTAAATTTCTGTGGGAGAATTCTTACATGTTTTCTTTGCTTTAAGTGTAACTGGCAGTTTTC

CATTGGTTTACCTGTGAAATAGTTCAAAGCCAAGTTTATATACAATTATATCAGTCCTCTTTCAAAGGTA
GCCATCATGGATCTGGTAGGGGGAAAATGTGTATTTTATTACATCTTTCACATTGGCTATTTAAAGACAA
AGACAAATTCTGTTTCTTGAGAAGAGAATATTAGCTTTACTGTTTGTTATGGCTTAATGACACTAGCTAA
TATCAATAGAAGGATGTACATTTCCAAATTCACAAGTTGTGTTTGATATCCAAAGCTGAATACATTCTGC

5   TTTCATCTTGGTCACATACAATTATTTTTACAGTTCTCCCAAGGGAGTTAGGCTATTCACAACCACTCAT
TCAAAAGTTGAAATTAACCATAGATGTAGATAAACTCAGAAATTTAATTCATGTTTCTTAAATGGGCTAC
TTTGTCCTTTTTGTTATTAGGGTGGTATTTAGTCTATTAGCCACAAAATTGGGAAAGGAGTAGAAAAAGC
AGTAACTGACAACTTGAATAATACACCAGAGATAATATGAGAATCAGATCATTTCAAAACTCATTTCCTA

10   TGTAACTGCATTGAGAACTGCATATGTTTCGCTGATATATGTGTTTTCACATTTGCGAATGGTTCCATT
CTCTCTCCTGTACTTTTTCCAGACACTTTTTTGAGTGGATGATGTTTCGTGAAGTATACTGTATTTTTAC
CTTTTTCCTTCCTTATCACTGACACAAAAAGTAGATTAAGAGATGGGTTTGACAAGGTTCTTCCCTTTTA
CATACTGCTGTCTATGTGGCTGTATCTTGTTTTTCCACTACTGCTACCACAACTATATTATCATGCAAAT
GCTGTATTCTTCTTTGGTGGAGATAAAGATTTCTTGAGTTTTGTTTTAAAATTAAAGCTAAAGTATCTGT
ATTGCATTAAATATAATATGCACACAGTGCTTTCCGTGGCACTGCATACAATCTGAGGCCTCCTCTCTCA

15   GTTTTTATATAGATGGCGAGAACCTAAGTTTCAGTTGATTTTACAATTGAAATGACTAAAAAACAAAGAA
GACAACATTAAAACAATATTGTTTCTA

TCF19

20   >gi|11419132|ref|XM_004567.1| Homo sapiens transcription factor 19 (SC1)
(TCF19), mRNA (SEQ ID NO:57)
ATGCTGCCCTGCTTCCAACTGCTGCGCATAGGGGGCGGCAGGGGCGGTGATCTCTACACCTTCCACCCCC
CCGCCGGGGCTGGCTGCACCTATCGCTTGGGCCACAGGGCCGACCTGTGTGATGTGGCCCTGCGGCCCCA

25   GCAGGAGCCTGGCCTCATCTCTGGGATCCACGCCGAACTGCATGCCGAGCCCCGGGGTGATGACTGGAGG
GTCAGCCTGGAAGACCACAGCAGCCAAGGTACTTTGGTCAATAATGTCCGACTCCCAAGAGGTCACAGGC
TGGAATTGAGTGATGGAGACCTCCTGACCTTTGGCCCTGAAGGGCCCCCAGGAACCAGCCCCTCGGAGTT
CTACTTCATGTTCCAACAAGTACGAGTCAAGCCTCAGGACTTTGCTGCCATTACCATCCCACGGTCTAGG

30   GGAGAAGCCCGGGTTGGGGCTGGTTTCCGGCCTATGCTGCCCTCCCAGGGGGCTCCACAGCGGCCTCTCA
GCACCTTCTCCCCTGCCCCCAAGGCCACACTGATCCTAAACTCCATAGGCAGCCTCAGCAAGCTCCGGCC
CCAGCCCCTCACCTTCTCCCCTAGTTGGGGTGGACCAAAGAGCCTGCCTGTTCCCGCCCCACCTGGGGAA
GTGGGGACCACGCCTTCTGCTCCACCCCAACGCAATCGGAGGAAATCTGTTCACCGAGTGTTGGCGGAAC
TGGATGATGAGAGTGAGCCTCCTGAGAACCCGCCACCGGTCCTTATGGAGCCCAGGAAGAAACTCCGTGT

35   AGACAAAGCCCCACTGACTCCCACTGGAAATCGACGTGGCCGTCCTCGGAAGTACCCAGTGAGCGCTCCC
ATGGCTCCCCCTGCAGTTGGGGGCGGGGAGCCCTGTGCAGCTCCTTGTTGCTGCCTGCCCCAGGAAGAGA
CAGTGGCCTGGGTTCAGTGTGATGGCTGTGACGTCTGGTTCCATGTGGCCTGTGTTGGCTGCAGCATCCA
GGCTGCCAGGGAGGCCGACTTCCGATGCCCAGGGTGCCGGGCTGGCATTCAGACCTAAGGTCCACCGCCA
AGGCACCATCGGACACACCTGCCCATGAGTAG

40   >gi|11419133|ref|XP_004567.1| transcription factor 19 (SC1) [Homo sapiens]
(SEQ ID NO:58)
MLPCFQLLRIGGGRGGDLYTFHPPAGAGCTYRLGHRADLCDVALRPQQEPGLISGIHAELHAEPRGDDWR
VSLEDHSSQGTLVNNVRLPRGHRLELSDGDLLTFGPEGPPGTSPSEFYFMFQQVRVKPQDFAAITIPRSR
GEARVGAGFRPMLPSQGAPQRPLSTFSPAPKATLILNSIGSLSKLRPQPLTFSPSWGGPKSLPVPAPPGE

45   VGTTPSAPPQRNRRKSVHRVLAELDDESEPPENPPPVLMEPRKKLRVDKAPLTPTGNRRGRPRKYPVSAP
MAPPAVGGGEPCAAPCCCLPQEETVAWVQCDGCDVWFHVACVGCSIQAAREADFRCPGCRAGIQT

>gi|15304334|ref|XM_004567.4| Homo sapiens transcription factor 19 (SC1)
(TCF19), mRNA (SEQ ID NO:59)

50   ATGCTGCCCTGCTTCCAACTGCTGCGCATAGGGGGCGGCAGGGGCGGTGATCTCTACACCTTCCACCCCC
CCGCCGGGGCTGGCTGCACCTATCGCTTGGGCCACAGGGCCGACCTGTGTGATGTGGCCCTGCGGCCCCA
GCAGGAGCCTGGCCTCATCTCTGGGATCCACGCCGAACTGCATGCCGAGCCCCGGGGTGATGACTGGAGG
GTCAGCCTGGAAGACCACAGCAGCCAAGGTACTTTGGTCAATAATGTCCGACTCCCAAGAGGTCACAGGC
TGGAATTGAGTGATGGAGACCTCCTGACCTTTGGCCCTGAAGGGCCCCCAGGAACCAGCCCCTCGGAGTT

55   CTACTTCATGTTCCAACAAGTACGAGTCAAGCCTCAGGACTTTGCTGCCATTACCATCCCACGGTCTAGG
GGAGAAGCCCGGGTTGGGGCTGGTTTCCGGCCTATGCTGCCCTCCCAGGGGGCTCCACAGCGGCCTCTCA
GCACCTTCTCCCCTGCCCCCAAGGCCACACTGATCCTAAACTCCATAGGCAGCCTCAGCAAGCTCCGGCC
CCAGCCCCTCACCTTCTCCCCTAGTTGGGGTGGACCAAAGAGCCTGCCTGTTCCCGCCCCACCTGGGGAA
GTGGGGACCACGCCTTCTGCTCCACCCCAACGCAATCGGAGGAAATCTGTTCACCGAGTGTTGGCGGAAC

60   TGGATGATGAGAGTGAGCCTCCTGAGAACCCGCCACCGGTCCTTATGGAGCCCAGGAAGAAACTCCGTGT
AGACAAAGCCCCACTGACTCCCACTGGNNNNNNNNNNNNNNNNCGTCCTCGGAAGTACCCAGTGAGCGCTCCC

```
ATGGCTCCCCCTGCAGTTGGGGGCGGGGAGCCCTGTGCAGCTCCTTGTTGCTGCCTGCCCCAGGAAGAGA
CAGTGGCCTGGGTTCAGTGTGATGGCTGTGACGTCTGGTTCCATGTGGCCTGTGTTGGCTGCAGCATCCA
GGCTGCCAGGGAGGCCGACTTCCGATGCCCAGGGTGCCGGGCTGGCATTCAGACCTAAGGTCCACCGCCA
AGGCACCATCGGACACACCTGCCCATGAGTAG
```
5

HFB101S

>gi|7706506|ref|NM_016325.1| Homo sapiens zinc finger protein 274 (ZNF274),
10 mRNA (SEQ ID NO:60)
```
GTCCAGCGCCTTTCCTTTCTCCAGCTTCTGCTCTGCCCCAAGAGTGGTCGCCTTCGTGGCAGGGCACGAC
TCTCTCCCAGCTCGGGTCGCCGCCCACATTAGTGTGGGGCCCTGCGGCCTAGCGTCCCTCACCAGAGGCC
TCCCCTTGCCTAGCTGGACCGCCGAGGGACATCGACGAGTATCCTCCTCCTGCTGTCCCCGGCTTCGCCT
GCCGCCCCTAACCGGCCAGTCAAGATGGCCGCCGCTGGGTGAGGCAAGCTGGCGCGCCGCGGGGCGTCT
GGGAGTTGTAGTTCGGGACGGCGGGCTGACGCACTTCGCCGCCGGCCGACGGGCGCCATTGTGCGGCGCG
CGCCGGGACTCTGCCCACTTCCACCAGAGACACATTGAGAAGGAGGAAACTATGGCCTCCAGGCTTCCGA
CGGCCTGGTCCTGTGAACCAGTGACCTTTGAAGATGTAACACTGGGTTTTACCCCGGAAGAGTGGGGACT
GCTGGACCTCAAACAGAAGTCCCTGTACAGGGAAGTGATGCTGGAGAACTACAGGAACCTGGTCTCAGTG
GAGTATCCTGAGCTCCAGCTGGACCCTAAATTGGATCCTCTTCCTGCTGAGAGTCCCCTAATGAACATTG
AGGTTGTTGAGGTCCTCACACTGAACCAGGAGGTGGCTGGTCCCCGGAATGCCCAGATCCAGGCCCTATA
TGCTGAAGATGGAAGCCTGAGTGCAGATGCCCCCAGTGAGCAGGTCCAACAGCAGGGCAAGCATCCAGGT
GACCCTGAGGCCGCGCGCCAGAGGTTCCGGCAGTTCCGTTATAAGGACATGACAGGTCCCCGGGAGGCCC
TGGACCAGCTCCGAGAGCTGTGTCACCAGTGGCTACAGCCTAAGGCACGCTCCAAGGAGCAGATCCTGGA
GCTGCTGGTGCTGGAGCAGTTCCTAGGTGCACTGCCTGTGAAGCTCCGGACATGGGTGGAATCGCAGCAC
CCAGAGAACTGCCAAGAGGTGGTGGCCCTGGTAGAGGGTGTGACCTGGATGTCTGAGGAGGAAGTACTTC
CTGCAGGACAACCTGCCGAGGGCACCACCTGCTGCCTCGAGGTCACTGCCCAGCAGGAGGAGAAGCAGGA
GGATGCAGCCATCTGCCCAGTGACAGTGCTCCCTGAGGAGCCAGTGACCTTCCAGGATGTGGCTGTGGAC
TTCAGCCGGGAGGAGTGGGGGCTGCTGGGCCCGACACAGAGGACCGAGTACCGCGATGTGATGCTGGAGA
CCTTTGGGCACCTGGTCTCTGTGGGGTGGGAGACTACACTGGAAAATAAAGAGTTAGCTCCAAATTCTGA
CATTCCTGAGGAAGAACCAGCCCCAGCCTGAAAGTACAAGAATCCTCAAGGGATTGTGCCTTGTCCTCT
ACATTAGAAGATACCTTGCAGGGTGGGGTCCAGGAAGTCCAAGACACAGTGTTGAAGCAGATGGAGTCTG
CTCAGGAAAAAGACCTTCCTCAGAAGAAGCACTTTGACAACCGTGAGTCCCAGGCAAACAGTGGTGCTCT
TGACACAAACCAAGTTTCGCTCCAGAAAATTGACAACCCTGAGTCCCAGGCAAACAGTGGCGCTCTTGAC
ACAAACCAAGTTTTGCTCCACAAAATTCCTCCTAGAAAACGATTGCGCAAACGTGACTCACAAGTTAAAA
GTATGAAACATAATTCACGTGTAAAAATTCATCAGAAGAGCTGTGAAAGGCAAAAGGCCAAGGAAGGCAA
TGGTTGTAGGAAAACCTTCAGTCGGAGTACTAAACAGATTACGTTTATAAGAATTCACAAGGGGAGCCAA
GTTTGCCGATGCAGTGAATGTGGTAAAATATTCCGGAACCCAAGATACTTTTCTGTGCATAAGAAATCC
ATACCGGAGAGAGGCCCTATGTGTGTCAAGACTGTGGGAAAGGATTTGTTCAGAGCTCTTCCCTCACACA
GCATCAGAGAGTTCATTCTGGAGAGAGACCATTTGAATGTCAGGAGTGTGGGAGGACCTTCAATGATCGC
TCAGCCATCTCCCAGCACCTGAGGACTCACACTGGCGCTAAGCCCTACAAGTGTCAGGACTGTGGAAAAG
CCTTCCGCCAGAGCTCCCACCTCATCAGACATCAGAGGACTCACACCGGGGAGCGCCCATATGCATGCAA
CAAATGTGGAAAGGCCTTCACCCAGAGCTCACACCTTATTGGGCACCAGAGAACCCACAATAGGACAAAG
CGAAAGAAGAAACAGCCTACCTCATAGCTCTCAAGCCAGTTGAAGAAACCTTGCCTTTTCAGCTTGACCC
TGCAATATAACATGCACAGGCCTGCTTGTAATCAGGACTGAATGTGAAAGGGAAGTATTGAGTGAGGAC
ATTCCCAAAACCAAAGGACAACTGAGGAGACTGCCCAGCACATAATGAATAAATAAGAAAATGAGTGAGG
AGTTATTAACATCATTTGGAAAAAAGATTTCCCATTCACTTGATATTGTTTGTTCACTCATTTAGTCATT
AAAAGTGAGATTAATAAAATCTGAAAATGTTATATAATAACTTTAAAAAGCCAGGTAATTAATAATCTGC
ACTGATATTACATCCACAGTACCACAGTATTTATGTGTATGAATTAAGGATTAAAAGATAATGTGGATAA
ATAAACTATTGATCTATG
```
50

>gi|7706507|ref|NP_057409.1| zinc finger protein 274, isoform a; KRAB zinc
finger protein HFB101; zinc finger protein zfp2 [Homo sapiens] (SEQ ID
NO:61)
```
MASRLPTAWSCEPVTFEDVTLGFTPEEWGLLDLKQKSLYREVMLENYRNLVSVEYPELQLDPKLDPLPAE
SPLMNIEVVEVLTLNQEVAGPRNAQIQALYAEDGSLSADAPSEQVQQQGKHPGDPEAARQRFRQFRYKDM
TGPREALDQLRELCHQWLQPKARSKEQILELLVLEQFLGALPVKLRTWVESQHPENCQEVVALVEGVTWM
SEEEVLPAGQPAEGTTCCLEVTAQQEEKQEDAAICPVTVLPEEPVTFQDVAVDFSREEWGLLGPTQRTEY
RDVMLETFGHLVSVGWETTLENKELAPNSDIPEEEPAPSLKVQESSRDCALSSTLEDTLQGGVQEVQDTV
LKQMESAQEKDLPQKKHFDNRESQANSGALDTNQVSLQKIDNPESQANSGALDTNQVLLHKIPPRKRLRK
RDSQVKSMKHNSRVKIHQKSCERQKAKEGNGCRKTFSRSTKQITFIRIHKGSQVCRCSECGKIFRNPRYF
SVHKKIHTGERPYVCQDCGKGFVQSSSLTQHQRVHSGERPFECQECGRTFNDRSAISQHLRTHTGAKPYK
```

CQDCGKAFRQSSHLIRHQRTHTGERPYACNKCGKAFTQSSHLIGHQRTHNRTKRKKKQPTS

RERE

>gi|8096339|dbj|AB036737.1| Homo sapiens mRNA for RERE, complete cds (SEQ
ID NO:62)
CGAAAATCACCATCATCCTCAGGGAGCCCCCTCGCCCTCGGAAAACAAAACCAGACATGATCCTGGGCTC
CTGCTTCTGAAAGATCTGCCCGACCGAGGAGGAGGGAGCCTAGGTTCAAGTCGCAGGAAAGGTCACGGCA
GGGAAGCCCCCTCAGCCCCTCCGCGTCGCCTCCCCGCGGGCCGCCCCTCTCGGCCCTGCGCCCCGCGGCC
CCGCAGCCTCCGCGCGGCCTCCCGCCCCATCCCCGCCACTTTCCCCGGAGCTGGGCCGCGAACACCTCAC
TATTGATGGTGTAGCGCTTTAGGGGAAGGAGGTGATTGTCTAGGAGTAACTGTATGTGCGCTCAGCACGG
GTCACAACCACCGTGGGATCCACTTACAGAGAAGATTTGTGGAGAGCTTCAGCCTGTTTATGGTGTTACT
GAGAAAGGAATTGTGGCCATTTGATATAATTCTGGATGGAAGACCTCTTATTTGTATTACATTGAATTAG
AGCATTTTTGAAACGTTGGGGTTGTTGGAGTGGTTGGATTTTCCCTGGAATTGAGTGAGAAATTCAGAAG
ACTGAAGCCCAGGCTCACTGTCTACCTTTCACGGAGGCCTAGCCGTGAGAGGACAGAAGAAGGCACGTGG
CGAATC<u>ATG</u>ACAGCGGACAAAGACAAAGACAAAGACAAAGAGAAGGACCGGGACCGAGACCGGGACCGAG
AGAGAGAGAAAAGAGACAAAGCAAGAGAGAGTGAGAATTCAAGGCCACGCCGGAGCTGTACCTTGGAAGG
AGGAGCCAAAAATTATGCTGAGAGTGATCACAGTGAAGACGAGGACAATGACAACAATGGAGCCACCACG
GAGGAGTCCACGAAGAAGAATAAGAAGAAACCACCGAAAAAAAAGTCTCGTTATGAAAGGACAGATACCG
GTGAGATAACATCCTACATCACTGAAGATGATGTGGTCTACAGACCAGGAGACTGTGTGTATATCGTGTG
TCGGAGGCCAAACACACCGTATTTCATCTGTAGCATTCAAGACTTCAAACTGGTCCACAACTCCCAGGCC
TGTTGCAGATCTCCAACTCCTGCTTTGTGTGACCCCCCAGCATGCTCTCTGCCGGTGGCATCACAGCCAC
CGCAGCATCTTTCTGAAGCCGGGAGAGGGCCTGTAGGGAGTAAGAGGGACCATCTCCTCATGAACGTCAA
ATGGTACTACCGTCAATCTGAGGTTCCAGATTCTGTGTATCAGCATTTGGTTCAGGATCGACATAATGAA
AATGACTCTGGAAGAGAACTTGTCATTACAGACCCAGTTATCAAGAACCGAGAGCTCTTCATTTCTGATT
ACGTTGACACTTACCATGCTGCTGCCCTTAGAGGGAAGTGTAACATCTCCCATTTTTCTGACATATTTGC
TGCTAGAGAGTTTAAAGCCCGAGTGGATTCATTTTTCTACATATTAGGATATAACCCTGAGACAAGGAGG
CTGAACAGTACCCAGGGGGAGATTCGTGTCGGTCCTAGTCATCAGGCCAAACTTCCAGATCTGCAACCAT
TTCCTTCTCCAGATGGTGATACAGTGACCCAACATGAGGAACTGGTCTGGATGCCTGGAGTTAACGACTG
TGACCTCCTTATGTACTTGAGGGCAGCAAGGAGCATGGCGGCATTTGCAGGAATGTGTGATGGAGGCTCT
ACAGAGGACGGCTGTGTCGCAGCCTCTCGGGATGACACCACTCTGAATGCACTGAACACACTGCATGAAA
GCGGTTACGATGCTGGCAAAGCCCTGCAGCGCCTGGTGAAGAAGCCTGTGCCCAAGCTCATCGAGAAGTG
CTGGACCGAGGACGAAGTGAAACGCTTCGTTAAGGGACTCAGGCAGTACGGGAAGAACTTCTTCAGAATT
AGAAAGGAGCTGCTTCCCAATAAGGAAACAGGGGAGCTGATCACCTTCTATTACTATTGGAAGAAGACCC
CCGAAGCAGCCAGCTCCCGAGCCCATCGTAGGCACCGCAGGCAGGCCGTGTTCAGGAGGATTAAGACTCG
CACCGCGTCCACACCCGTCAACACACCCTCCAGACCCCGTCCAGTGAATTCTTGGACCTAAGTTCAGCC
AGTGAAGATGACTTCGACAGTGAGGACAGTGAGCAGGAGCTGAAGGGGTACGCCTGCCGCCACTGCTTCA
CCACCACCTCCAAAGATTGGCACCACGGAGGCCGGGAGAACATCCTGCTTTGCACCGACTGTCGCATCCA
CTTCAAGAAATACGGTGAGCTCCCGCCCATTGAGAAGCCCGTGGACCCGCCACCGTTTATGTTCAAACCC
GTCAAGGAAGAGGATGATGGGCTCAGTGGGAAGCATAGCATGAGGACACGGCGGAGTCGGGGCTCGATGT
CGACACTACGCAGTGGTCGGAAGAAGCAGCCAGCCAGCCCTGATGGTCGCACCTCACCCATCAATGAAGA
CATCCGCTCCAGCGGCCGGAACTCCCCCAGCGCTGCCAGTACCTCCAGCAATGACAGTAAAGCAGAGACA
GTGAAGAAGTCGGCCAAGAAGGTGAAGGAGGAAGCCTCTTCCCTCTTAAGAGTAACAAACGCCAGCGGG
AGAAGGTGGCCTCTGATACGGAGGAGGCTGACAGGACCAGCTCCAAGAAGACAAAAACGCAGGAGATCAG
CAGGCCCAACTCGCCATCTGAAGGTGAGGGAGAGAGTTCAGACAGTCGCAGCGTCAACGATGAGGGTAGC
AGTGACCCCAAAGACATCGACCAGGACAATCGCAGCACGTCCCCGAGCATCCCCAGCCCCCAGGACAATG
AGAGTGACTCGGACTCGTCAGCCCAGCAGCAGATGCTGCAGGCCCAGCCCCCAGCCTTGCAGGCTCCCAC
TGGGGTCACCCCAGCTCCCTCCTCAGCTCCTCCAGGGACCCCTCAGCTGCCCACGCCAGGGCCCACGCCC
TCTGCCACTGCAGTTCCCCCACAGGGCTCCCCCACGGCCTCCCAGGCCCCTAACCAGCCGCAGGCTCCCA
CAGCGCCTGTTCCCCACACCCACATCCAACAGGCACCGGCCTTGCACCCCCAGCGGCCGCCCTCACCGCA
TCCCCCGCCGCATCCCTCGCCACATCCCCCACTGCAGCCTCTGACTGGGTCGGCGGGCCAGCCTTCTGCA
CCCTCTCATGCCCAGCCCCCACTGCACGGTCAGGGCCCACCCGGCCCTCACAGCCTGCAGGCTGGGCCCC
TGCTGCAGCACCCAGGCCCCCACAGCCCTTTGGCCTCCCTCCCCAGGCCTCCCAAGGCCAGGCCCCTCT
GGGGACCTCCCCAGCAGCAGCGTACCCTCACACCTCCCTGCAGCTGCCAGCCTCTCAGTCAGCGCTGCAG
TCCCAACAGCCTCCACGGGAGCAGCCCCTGCCACCAGGGCCCTTGGCCATGCCCCACATCAAGCCCCCGC
CTACCACTCCCATCCCCCAGCTGGCGGCGCCACAGGCCCACAAGCACCCTCCCCACCTCTCGGGGCCCTC
ACCCTTCTCCATGAATGCCAACCTGCCTCCCCCTCCAGCCCTGAAGCCCCTGAGCTCCCTGTCCACACAT
CACCCCCCGTCGGCTCACCCCCCACCCCTGCAACTCATGCCTCAGAGCCAGCCATTGCCCTCCTCGCCCG
CCCAGCCCCCCGGGCTGACCCAGAGCCAGAACCTGCCCCCGCCCCCTGCCTCCCACCCCCCTACAGGCCT
CCACCAGGTGGCCCCCCAACCCCCGTTTGCTCAGCACCCCTTTGTCCCTGGAGGCCCTCCTCCCATCACC

CCTCCGACCTGCCCCTCCACCTCTACCCCACCGGCGGGACCTGGCACCTCGGCCCAGCCACCCTGCTCTG
GTGCGGCGGCTTCAGGAGGCAGCATAGCGGGGGGGTCGTCCTGCCCACTCCCCACCGTCCAGATCAAGGA
GGAGGCTCTGGACGACGCTGAGGAGCCTGAGAGCCCCCCTCCCCCACCAAGGAGCCCGTCCCCGGAGCCC
ACTGTGGTGGACACCCCCAGTCACGCCAGCCAGTCAGCTAGGTTCTACAAACACCTGGACCGGGGCTACA
5 ACTCGTGTGCCCGGACAGACCTGTACTTCATGCCTCTGGCCGGGTCCAAGCTGGCCAAGAAGAGGGAGGA
GGCCATTGAGAAGGCCAAGCGCGAGGCTGAGCAGAAAGCCCGAGAGGAGCGAGAGCGGGAGAAGGAGAAG
GAGAAGGAGCGGGAGCGGGAGCGAGAGCGGGAGCGCGAGGCAGAGCGGGCGGCTAAGGCGTCCAGCTCAG
CGCATGAAGGTCGCCTCAGTGACCCACAGCTCAGTGGTCCTGGCCACATGCGGCCATCCTTCGAGCCACC
ACCAACCACCATTGCTGCTGTGCCCCCCTACATCGGGCCCGACACACCTGCCCTTCGGACTCTGAGCGAG
10 TACGCCCGGCCCCACGTCATGTCGCCCACCAACCGCAACCACCCCTTCTACATGCCCCTTAACCCCACGG
ACCCCCTGCTGGCCTACCACATGCCTGGCCTCTACAACGTCGACCCCACCATCCGCGAGCGGGAGCTCCG
GGAGCGGGAGATCCGAGAGCGGGAGATCCGAGAGCGGGAGCTGCGGGAGAGGATGAAGCCGGGCTTCGAG
GTGAAGCCCCAGAGCTGGACCCCCTGCACCCAGCCGCCAACCCCATGGAGCACTTTGCCCGGCACAGCG
CCCTCACCATCCCCCCGACCGCCGGGCCCCACCCTTTTGCTTCTTTCCACCCGGGCCTGAACCCCTTGGA
15 GAGGGAGAGACTGGCCCTGGCGGGCCCCCAGCTGCGGCCCGAGATGAGCTACCCTGACAGACTGGCAGCC
GAGCGTATCCACGCAGAGCGCATGGCATCGCTGACCAGCGATCCCCTGGCCCGACTGCAGATGTTCAACG
TGACTCCGCACCATCACCAGCACTCTCACATTCACTCCCACCTCCACCTCCACCAGCAGGACCCCCTCCA
CCAAGGTTCAGCAGGCCCCGTTCACCCGCTGGTCGACCCCCTGACTGCCGGTCCCCACCTGGCTCGCTTC
CCCTACCCGCCTGGCACTCTCCCCAACCCTCTGCTTGGACAGCCCCCACACGAGCACGAGATCGTTCGCC
20 ACCCAGTTTTCGGCACCCCCTACCCCCGTGACCTGCCTGGGGCCATCCCCACCCCCCATGTCAGCAGCCCA
CCAGCTGCAGGCCATGCATGCCCAGTCGGCCGAGCTGCAGAGACTGGCCATGGAGCAGCAGTGGCTGCAT
GGACACCCCCACATGCATGGTGGCCACCTACCAAGTCAGGAAGATTATTACAGTCGACTGAAGAAAGAAG
GTGACAAGCAGTTATAAGTTATTTATTTGTTAACGCTGGCTGTGGAAACCCCAGTTCTTGGGGGAGAAAC
AGGACTTTTTACATAAAATAGGAGCTGCAAAAGCAAAAGAATATCTTCTAAAGATTTCTTTATATATTT
25 AAAAACCCACAACTAAAAATGTATCCACATAGTAGTGTTCGTTTGTCGAGAGGATTTCCTGAGACTGGTT
TGGATCTCCCTGCATGACAGTCCCCCAGAAACTTAGTGAGTCCTGGACTGGACTGAACATCCAGAAAGCT
TCCCTGCAATCTTGGGGTTTGGCTTTAGTTTTCTTTTCCTTGATTTCTCAGTAGGTGCTAGAATCCAGTT
CACACCCTTCACTGTGCGTGCAGACACACTGACACACTCCGCCACGAGTGCTCCAGAGCCCACGAGGCTT
GCAGATCGGGGGCATAGGAATTTGGAATCCAAGAGCTATAATTTTTAAAAAAAAAAAATCTTTTATTTTAA
30 TACATTGTAGGAAATCTTCATAATTGGAGAAAGTTCTGCAGCATGGCTTTTTACGTCTGTAAATAAATAA
TTTTAGAACAGCCTTTTTTTCTTCCATAAACTACTATTGTGATCTATTTTTCCAGCCATGTCACTAATT
GTGAATTCCTACCAACTATTGACAGAATACAGAGTTGATTTTTTAATAAAAAGTTATATATAATTATCCC
TTTAATTAAAGGGAGCAAAGGGGCGTTTCACATGGACAGAGGCTTGGACCGAGGCCTGGTCACAGCAGCG
AGCATCCAGGGTTTGCAGGACGATGTTACAGACTCTGTTTTCTGCCTGGCGTTTCACTTGTGTCTGCTC
35 CTAGCCTGTGCTCTGCCAGCAGCACAGACATCTGCTCCATCAGACCTCTTCCATTTTGCACAGGGAGTGC
AGGAGGTGAATGTTCACTTTCTGTTCTCCAGTGTCACTGTTCTGTTTCCACGGGATGGAAAGCGCATGGG
CCTGTGTCCATTGTAGATTTCCTTCTAGATTTCTGTGTACACACACTTGATTGTTCTGGATGAATGTCTT
TTTTAATACTCCGAAAATTTCATCATCTAAGAAAATGATTCCATACAAATAACTCAGCACACAAGTGACC
CAGGACATATGCCTGCCAAAGGGATGTGTTAGAAGGCTGCCTTCTCATGCGCATTGTCACTTGGATCTTG
40 TGGTGAGGACGGCCCCATCTTTCTTGCCACAGATTGAGGCCACTTTTGAGCAAGGGAGATCCTGGAGTTA
AGACAGGTGTTGGGGGCAGCCTGTATTTTACCCTAGGGGCAGGTCTGCATGGTGACCCCACATTGCACTG
GTAAACCATTTGAGTCCCACTCTTCATCCTGGAAGTGGGAACTGGAGTCCCACCCACAGTGCATTCAGAA
AGCATGCTGTGTGGGGGCTGCTTCTCAGGAGGCCAGGCCCTTCTGAGCGGAACCGTCCTGGAGAGAGCCT
GCCCTCGTTTCCAGGCTGCAGCCGTAACGCACTTTCTCCCAGGCTGAGGGCGGGTGTTCTGGGGTGTCTG
45 CCCTCTGTCGGCCCTGCTTCCTGCCAGGACGTGGCCTCTTCCGATCCTTTTCTCTCAGACACTGGAGGTC
TCTTCTGCCATTGTGCTGGTCCATCCCAAGAATTGTAGGACAGAGACCACACTGGGTCGGCGGACACAA
AGTCCATCCAGGACCCAGGCCGCAGAGGGAGCAGGAAGAGTTGCTGATAGTTTGATCTAGAAACCAGCAG
CTACTGGCTCAAATTCAGGTTCTGGCGTCAAATAGCGACATTTCCAGTTTCTCTTAAAAACCGTGTTTGG
TTTCAGTTGGGATAGGCTTGTTTTGTCTGTTGAAAATGTTTCTAGTTTTTTTTCTTTCATTTTTCTCTCA
50 TTCCATTTCTGCCTTAACTTTAGTTTGTTCACAGGGAGGCAAAGCTGACATGAACCTTTTGTCGTGGGAC
TTCAGGCCACATTGGCTTGAAGGCATTCGTTTCCTTCTGGGGTGGGGACAGGCCCTCATGGCAGGCTTGT
TCCCGTGGCTCTGAGCGAGGCCTCTTCCTGCTGGGCTCCCAGACTCCTGCATCCAGGCCCCCACCTTCTC
GGCTTCTGGTTTTTCTTTCTTTTTGGTAGAACACAACATCTACCATTCAGTTAAACCTTCTTTATCCCCT
CCTTTGGCATCCATTTTTCCAAAGAAGAGTCGAGTCCTCTGAGGTCTGTGCTTGAAAACCGTCCGAAGGC
55 ATTCTTGTTAGCTTTGCTTTTCTCCCCATATCCCAAGGCGAACGGCTGAGATTCTTCCATCTAAAAAACC
CTCGACCCGAAACCCTCACCAGATAAACTACAGTTTGTTTAGGAGGCCCTGACCTTCATGGTGTCTTTGA
AGCCCAACCACTCGGTTTCCTTCGGATTTTCCTCCCTTTGTTCGGGGTTTGGTTTGGCTCCTCTGTGTGT
GTCCGTATCTTGTTCGGTGTCCTCGAGGTTGAGCTTCACTCCACTGCGGCAGAGGCAGCGTGCACACTCG
GATTTGCTACGTTTCTATATATCTTGAAGCTAAATGTATATATGAGTAGTTTGCCATGAGATAACACAGT
60 GTAAACAGTAGACACCCAGAAATCGTGACTTCTGTGTTCTCTCCATTTGAGTATTTTGTAATTTTTTTGA
AATATTTGTGGACATAAATAAAACCAAGCTACACTACACCCTTGGAAAAAAAAA

>gi|8096340|dbj|BAA95898.1| RERE [Homo sapiens] (SEQ ID NO:63)
MTADKDKDKDKEKDRDRDRDREREKRDKARESENSRPRRSCTLEGGAKNYAESDHSEDEDNDNNGATTEE
STKKNKKKPPKKKSRYERTDTGEITSYITEDDVVYRPGDCVYIVCRRPNTPYFICSIQDFKLVHNSQACC

5   RSPTPALCDPPACSLPVASQPPQHLSEAGRGPVGSKRDHLLMNVKWYYRQSEVPDSVYQHLVQDRHNEND
SGRELVITDPVIKNRELFISDYVDTYHAAALRGKCNISHFSDIFAAREFKARVDSFFYILGYNPETRRLN
STQGEIRVGPSHQAKLPDLQPFPSPDGDTVTQHEELVWMPGVNDCDLLMYLRAARSMAAFAGMCDGGSTE
DGCVAASRDDTTLNALNTLHESGYDAGKALQRLVKKPVPKLIEKCWTEDEVKRFVKGLRQYGKNFFRIRK
ELLPNKETGELITFYYYWKKTPEAASSRAHRRHRRQAVFRRIKTRTASTPVNTPSRPPSSEFLDLSSASE

10 DDFDSEDSEQELKGYACRHCFTTTSKDWHHGGRENILLCTDCRIHFKKYGELPPIEKPVDPPPFMFKPVK
EEDDGLSGKHSMRTRRSRGSMSTLRSGRKKQPASPDGRTSPINEDIRSSGRNSPSAASTSSNDSKAETVK
KSAKKVKEEASSPLKSNKRQREKVASDTEEADRTSSKKTKTQEISRPNSPSEGEGESSDSRSVNDEGSSD
PKDIDQDNRSTSPSIPSPQDNESDSDSSAQQQMLQAQPPALQAPTGVTPAPSSAPPGTPQLPTPGPTPSA
TAVPPQGSPTASQAPNQPQAPTAPVPHTHIQQAPALHPQRPPSPHPPPHPSPHPPLQPLTGSAGQPSAPS

15 HAQPPLHGQGPPGPHSLQAGPLLQHPGPPQPFGLPPQASQGQAPLGTSPAAAYPHTSLQLPASQSALQSQ
QPPREQPLPPGPLAMPHIKPPPTTPIPQLAAPQAHKHPPHLSGPSPFSMNANLPPPPALKPLSSLSTHHP
PSAHPPPLQLMPQSQPLPSSPAQPPGLTQSQNLPPPPASHPPTGLHQVAPQPPFAQHPFVPGGPPPITPP
TCPSTSTPPAGPGTSAQPPCSGAAASGGSIAGGSSCPLPTVQIKEEALDDAEEPESPPPPPRSPSPEPTV
VDTPSHASQSARFYKHLDRGYNSCARTDLYFMPLAGSKLAKKREEAIEKAKREAEQKAREEREREKEKEK

20 ERERERERERERAAERAAKASSSAHEGRLSDPQLSGPGHMRPSFEPPPTTIAAVPPYIGPDTPALRTLSEYA
RPHVMSPTNRNHPFYMPLNPTDPLLAYHMPGLYNVDPTIRERELREREIREREIRERELRERMKPGFEVK
PPELDPLHPAANPMEHFARHSALTIPPTAGPHPFASFHPGLNPLERERLALAGPQLRPEMSYPDRLAAER
IHAERMASLTSDPLARLQMFNVTPHHHQHSHIHSHLHLHQQDPLHQGSAGPVHPLVDPLTAGPHLARFPY
P̂PGTLPNPLLGQPPHEHEIVRHPVFGTPYPRDLPGAIPPPMSAAHQLQAMHAQSAELQRLAMEQQWLHGH

25 PHMHGGHLPSQEDYYSRLKKEGDKQL

sudD

30 >gi|4507298|ref|NM_003831.1| Homo sapiens sudD (suppressor of bimD6,
Aspergillus nidulans) homolog (SUDD), mRNA (SEQ ID NO:64)
GAATTCGCGGCCGCCCCGCCTGTGTCCTCGGCGGAGCCTGCTGCCCGTCCTGCCACCTCTCTGCTCTGTT
CTTGTCTCTGCCTTCATTCCCGAATGGATCTGGTAGGAGTGGCATCGCCTGAGCCCGGGACGGCAGCGGC
CTGGGGACCCAGCAAGTGTCCATGGGCTATTCCTCAAAATACAATATCTTGTTCTTTGGCTGATGTAATG

35 AGTGAACAGCTGGCCAAAGAATTGCAGTTAGAAGAAGAAGCTGCCGTTTTTCCTGAAGTTGCTGTTGCTG
AAGGACCATTTATTACTGGAGAAAACATTGATACTTCCAGTGACCTTATGCTGGCTCAGATGCTACAGAT
GGAATATGACAGAGAATATGATGCACAGCTTAGGCGTGAAGAAAAAAATTCAATGGAGATAGCAAAGTT
TCCATTTCCTTTGAAAATTATCGAAAGTGCATCCTTATGAAGACAGCGATAGCTCTGAAGATGAGGTTG
ACTGGCAGGATACTCGTGATGATCCCTACAGACCAGCAAAACCGGTTCCCACTCCTAAAAAGGGCTTTAT

40 TGGAAAAGGAAAGATATCACCACCAAACATGATGAAGTAGTATGTGGGAGAAAGAACACAGCAAGAATG
GAAAATTTTGCACCTGAGTTTCAGGTAGGAGATGGAATTGGAATGGATTTAAAACTATCAAACCATGTTT
TCAATGCTTTAAAACAACATGCCTACTCAGAAGAACGTCGAAGTGCCCGCCTACATGAGAAAAAGGAGCA
TTCTACAGCAGAAAAAGCAGTTGATCCTAAGACACGTTTACTTATGTATAAAATGGTCAACTCTGGAATG
TTGGAGACAATCACTGGCTGTATTAGTACAGGAAAGGAGTCTGTTGTCTTTCATGCATATGGAGGGAGCA

45 TGGAGGATGAAAAGGAAGATAGTAAAGTTATACCTACAGAATGTGCCATCAAGGTATTTAAAACAACCCT
TAATGAATTTAAGAATCGTGACAAATATATTAAAGATGATTTCAGGTTTAAAGATCGCTTCAGTAAACTA
AATCCACGTAAGATCCACCGCATGTGGGCAGAAAAAGAAATGCACAATCTCGCAAGAATGCAGAGAGCTG
GAATTCCTTGTCCAACAGTTGTACTACTGAAGAAACACATTTTAGTTATGTCTTTTATTGGCCATGATCA
AGTTCCAGCCCCTAAATTAAAAGAAGTAAAGCTCAATAGTGAAGAAATGAAAGAAGCCTACTATCAAACT

50 CTTCATTTGATGCGGCAGTTATATCATGAATGTACGCTTGTCCATGCTGACCTCAGTGAGTATAACATGC
TGTGGCATGCTGGAAAGGTCTGGTTGATCGATGTCAGTCAGTCAGTAGAACCTACCCACCCTCACGGCCT
GGAGTTCTTGTTCCGGGACTGCAGGAATGTCTCGCAGTTTTTCCAGAAAGGAGGAGTCAAGGAAGCCCTT
AGTGAACGAGAACTCTTCAATGCTGTTTCAGGCTTAAACATCACAGCAGATAATGAAGCTGATTTTTTAG
CTGAGATAGAAGCTTTGGAGAAAATGAATGAAGATCACGTTCAGAAGAATGGAAGGAAAGCTGCTTCATT

55 TTTGAAAGATGATGGAGACCCACCACTACTATATGATGAATAGCACTAATACCCACTGCTTCAGTGTTAA
CACAGCAGTGATTGTCAGCTGCCAATAGCAAATGAAGTTATGGGTGACTTGAAATACCAAAACCTGAGGA
GTGGGCAATGGTGCTTCTGTGCTTTTCCCCCTTGTAACCCATGTGCCAGATGTGTGGAATTTTTAGCTCA
GCATTGAGAGAATAAAATGTCACTACCTCTCATCTTATGAACAGGATAATATAATTCTTTAACAGCTATA
GGTTATCTGGCTGAAGTAGACCTAATTTTATGTGACTTGTGGTGTAAAATGTCTTGATGATAATTTTTAA

60 AACTTGGGTAACACTTCCAAATATGGGAGGAAAGGACAGATGTGTTTACAAGGGAGGATTTTACAACATA
CTTGCTTTATTCACCTCCCTGTTTTGTGTTGCGTCTTTCCTTGAATATTTTATTGGCCCAGAGTTAGCCT

```
TTCTCAATTATGTTTCCAGACTGTGGCCGTGATTCTAAAGGAAAATGTGTGCTCTTTAGTGGGTAGAACA
AATGGAAATTTGGTTTCAGAATGGCTGACAGAAATCGACATAAGTCATGTAATTTTTGTTGATATATCAT
GAAAATGAACAGAATTCTTTTTCCATACTTATATCTAAGAAAAGGCATCATAGGTTTCTGAAAGAGATAA
CTATATAACAGCTTTTTAACTATCCAGTCAACTTTCAGCTTTTCTACATTTAGGTAAAATGGTTAGGATA
```
5
```
TAACTCATGGTGTGGCTAATCTACATTTATCAATAAAATGTAAATTATCTGAAAGGACAGAATATAAGAT
TTAACCATGTTTGACATATTTTAATTTAGTTAATGAAGCAAAATTCAGTTTATATTTCACTAGAACTGTG
TACTTGATTGATTTTCAGAGAAATATCACAAATTAGAAATATTAAATCTAAGGATGAAAGGTATATATAA
AACAATTTGGGGGCCAGGCACGATGGCTCAAACCTGTAATCCCAGCACTTTGGGAGACCAAGGCGGGTGG
ATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGGCAACATGGCGAAACCCTGTCTCTACTAAAAATACA
```
10
```
AAAATTAGCCGGGTGTGGTGGCACTTCTCTGTAATCTCAGCTTCTCAGGAGGCTGAGACAGGAGAATCGC
TTGAACCCGGGAGGCAGAGGTTGCAGTGAGCTGAGATCATGCGCGGCCGCGAATTC
```

>gi|4507299|ref|NP_003822.1| sudD suppressor of bimD6 homolog (A.
nidulans); human homolog of Aspergillus nidulans sudD gene product; sudD

15
(suppressor of bimD6, Aspergillus nidulans) homolog [Homo sapiens] (SEQ ID
NO:65)
```
MDLVGVASPEPGTAAAWGPSKCPWAIPQNTISCSLADVMSEQLAKELQLEEEAAVFPEVAVAEGPFITGE
NIDTSSDLMLAQMLQMEYDREYDAQLRREEKKFNGDSKVSISFENYRKVHPYEDSDSSEDEVDWQDTRDD
PYRPAKPVPTPKKGFIGKGKDITTKHDEVVCGRKNTARMENFAPEFQVGDGIGMDLKLSNHVFNALKQHA
```
20
```
YSEERRSARLHEKKEHSTAEKAVDPKTRLLMYKMVNSGMLETITGCISTGKESVVFHAYGGSMEDEKEDS
KVIPTECAIKVFKTTLNEFKNRDKYIKDDFRFKDRFSKLNPRKIHRMWAEKEMHNLARMQRAGIPCPTVV
LLKKHILVMSFIGHDQVPAPKLKEVKLNSEEMKEAYYQTLHLMRQLYHECTLVHADLSEYNMLWHAGKVW
LIDVSQSVEPTHPHGLEFLFRDCRNVSQFFQKGGVKEALSERELFNAVSGLNITADNEADFLAEIEALEK
MNEDHVQKNGRKAASFLKDDGDPPLLYDE
```
25

>gi|22325376|ref|NM_003831.2| Homo sapiens sudD suppressor of bimD6 homolog
(A. nidulans) (SUDD), transcript variant 1, mRNA (SEQ ID NO:66)
```
CGGACGCGGCCGCCGCCGTCGCCGCCATCTGTCACCTCCACTCCGGCATCAGCAGCCAGTCGCCCGTGTC
CCGCCTGTCTCCTCGGCGGAGCCTGCTGCCCGTCCTGCCACCTCTCTGCTCTGTTCTTGTCTCTGCCTTC
```
30
```
ATTCCCGAATGGATCTGGTAGGAGTGGCATCGCCTGAGCCCGGGACGGCAGCGGCCTGGGGACCCAGCAA
GTGTCCATGGGCTATTCCTCAAAATACAATATCTTGTTCTTTGGCTGATGTAATGAGTGAACAGCTGGCC
AAAGAATTGCAGTTAGAAGAAGAAGCTGCCGTTTTTCCTGAAGTTGCTGTTGCTGAAGGACCATTTATTA
CTGGAGAAAACATTGATACTTCCAGTGACCTTATGCTGGCTCAGATGCTACAGATGGAATATGACAGAGA
ATATGATGCACAGCTTAGGCGTGAAGAAAAAAAATTCAATGGAGATAGCAAAGTTTCCATTTCCTTTGAA
```
35
```
AATTATCGAAAAGTGCATCCTTATGAAGACAGCGATAGCTCTGAAGATGAGGTTGACTGGCAGGATACTC
GTGATGATCCCTACAGACCAGCAAAACCGGTTCCCACTCCTAAAAAGGGCTTTATTGGAAAAGGAAAAGA
TATCACCACCAAACATGATGAAGTAGTATGTGGGAGAAAGAACACAGCAAGAATGGAAAATTTTGCACCT
GAGTTTCAGGTAGGAGATGGAATTGGAATGGATTTAAAACTATCAAACCATGTTTTCAATGCTTTAAAAC
AACATGCCTACTCAGAAGAACGTCGAAGTGCCCGCCTACATGAGAAAAAGGAGCATTCTACAGCAGAAAA
```
40
```
AGCAGTTGATCCTAAGACACGTTTACTTATGTATAAAATGGTCAACTCTGGAATGTTGGAGACAATCACT
GGCTGTATTAGTACAGGAAAGGAGTCTGTTGTCTTTCATGCATATGGAGGGAGCATGGAGGATGAAAAGG
AAGATAGTAAAGTTATACCTACAGAATGTGCCATCAAGGTATTTAAAACAACCCTTAATGAATTTAAGAA
TCGTGACAAATATATTAAAGATGATTTCAGGTTTAAAGATCGCTTCAGTAAACTAAATCCACGTAAGATC
ATCCGCATGTGGGCAGAAAAAGAAATGCACAATCTCGCAAGAATGCAGAGAGCTGGAATTCCTTGTCCAA
```
45
```
CAGTTGTACTACTGAAGAAACACATTTTAGTTATGTCTTTTATTGGCCATGATCAAGTTCCAGCCCCTAA
ATTAAAAGAAGTAAAGCTCAATAGTGAAGAAATGAAAGAAGCCTACTATCAAACTCTTCATTTGATGCGG
CAGTTATATCATGAATGTACGCTTGTCCATGCTGACCTCAGTGAGTATAACATGCTGTGGCATGCTGGAA
AGGTCTGGTTGATCGATGTCAGTCAGTCAGTAGAACCTACCCACCCTCACGGCCTGGAGTTCTTGTTCCG
GGACTGCAGGAATGTCTCGCAGTTTTTCCAGAAAGGAGGAGTCAAGGAAGCCCTTAGTGAACGAGAACTC
```
50
```
TTCAATGCTGTTTCAGGCTTAAACATCACAGCAGATAATGAAGCTGATTTTTTAGCTGAGATAGAAGCTT
TGGAGAAAATGAATGAAGATCACGTTCAGAAGAATGGAAGGAAAGCTGCTTCATTTTTGAAAGATGATGG
AGACCCACCACTACTATATGATGAATAGCACTAATACCCACTGCTTCAGTGTTAACACAGCAGTGATTGT
CAGCTGCCAATAGCAAATGAAGTTATGGGTGACTTGAAATACCAAAACCTGAGGAGTGGGCAATGGTGCT
TCTGTGCTTTTCCCCCTTGTAACCCATGTGCCAGATGTGTGGAATTTTTAGCTCAGCATTGAGAGAATAA
```
55
```
AATGTCACTACCTCTCATCTTATGAACAGGATAATATAATTCTTTAACAGCTATAGGTTATCTGGCTGAA
GTAGACCTAATTTTATGTGACTTGTGGTGTAAAATGTCTTGATGATAATTTTTAAAACTTGGGTAACACT
TCCAAATATGGGAGGAAAGGACAGATGTGTTTACAAGGGAGGATTTTACAACATACTTGCTTTATTCACC
TCCCTGTTTTGTGTTGCGTCTTTCCTTGAATATTTTATTGGCCCAGAGTTAGCCTTTCTCAATTATGTTT
CCAGACTGTGGCCGTGATTCTAAAGGAAAATGTGTGCTCTTTAGTGGGTAGAACAAATGGAAATTTGGTT
```
60
```
TCAGAATGGCTGACAGAAATCGACATAAGTCATGTAATTTTTGTTGATATATCATGAAAATGAACAGAAT
TCTTTTTCCATACTTATATCTAAGAAAAGGCATCATAGGTTTCTGAAAGAGATAACTATATAACAGCTTT
```

```
   TTAACTATCCAGTCAACTTTCAGCTTTTCTACATTTAGGTAAAATGGTTAGGATATAACTCATGGTGTGG
   CTAATCTACATTTATCAATAAAATGTAAATTATCTGAAAGGACAGAATATAAGATTTAACCATGTTTGAC
   GTATTTTAATTTAGTTAATGAAGCAAAATTCAGTTTATATTTCACTAGAACTGTGTACTTGATTGATTTT
   CAGAGAAATATCACAAATTAGAAATATTAAATCTAAGGATGAAAGGTATATATAAAACAATTTGGGGGCC
5  AGGCACGATGGCTCAAACCTGTAATCCCAGCACTTTGGGAGACCAAGGCGGGTGGATCACTTGAGGTCAG
   GAGTTCAAGACCAGCCTGGGCAACATGGCGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCCGGGTG
   TGGTGGCACTTCTCTGTAATCTCAGCTTCTCAGGAGGCTGAGACAGGAGAATCGCTTGAACCCGGGAGGC
   AGAGGTTGCAGTGAGCTGAGATCATGCCACTGCACTCCGGCCTAGGTGACAGAGGGAAACTCCATCTCCA
   GGAAAAAAAAAAAAAAACCCAATTTGGATACCAAATTAATCAACTAATTTGAGCTATCTGGCCTTACTCT
10 TAGTAGTTTTTAGTACGTGCTGGACACCACTTTTAAAAAGCAATCACTGTGCTAGAAAGTATATTGGCT
   TTGTTAGGATTAAAGTTCATTAACTTCAATGTAATCATGCCTCCTATTACTGAAGTCAGATTGGAACCAC
   TAAAGATCCAAACTTTCTGTCTGGTAATAGAAAGTAAAAATCTAGACATCATTTACATTTGAGAAGCTGT
   TTTTAACATTATTTTAAAATGCCAAATATGTTCTTTCTAGAAAAATATTTATTTTTGTTTTTGTTGGATA
   GCTTTTAATTACATTTCAGAGAGGTGTAATTTTGGGTAGATGCTCATTACATTTTTGAAAGGTTTATGAT
15 TCCAAAATAAAGATTTATATGACTGGTGATACTGGCTTTACAGAAATTTCAGAGAACTAATTTTTAAAAT
   CTTTAGCATTTAAAACTTTTTTTGTTTTGTTTTCTGACATATTCTGACAAAGAGCAGCAAACCACTG
```

KU 70

>gi|4503840|ref|NM_001469.1| Homo sapiens thyroid autoantigen 70kD (Ku
antigen) (G22P1), mRNA (SEQ ID NO:67)
```
   GGGGAGCAGTAGCCAACATGTCAGGGTGGGAGTCATATTACAAAACCGAGGGCGATGAAGAAGCAGAGGA
   AGAACAAGAAGAGAACCTTGAAGCAAGTGGAGACTATAAATATTCAGGAAGAGATAGTTTGATTTTTTTG
25 GTTGATGCCTCCAAGGCTATGTTTGAATCTCAGAGTGAAGATGAGTTGACACCTTTTGACATGAGCATCC
   AGTGTATCCAAAGTGTGTACATCAGTAAGATCATAAGCAGTGATCGAGATCTCTTGGCTGTGGTGTTCTA
   TGGTACCGAGAAAGACAAAAATTCAGTGAATTTTAAAAATATTTACGTCTTACAGGAGCTGGATAATCCA
   GGTGCAAAACGAATTCTAGAGCTTGACCAGTTTAAGGGGCAGCAGGGACAAAAACGTTTCCAAGACATGA
30 TGGGCCACGGATCTGACTACTCACTCAGTGAAGTGCTGTGGGTCTGTGCCAACCTCTTTAGTGATGTCCA
   ATTCAAGATGAGTCATAAGAGGATCATGCTGTTCACCAATGAAGACAACCCCCATGGCAATGACAGTGCC
   AAAGCCAGCCGGGCCAGGACCAAAGCCGGTGATCTCCGAGATACAGGCATCTTCCTTGACTTGATGCACC
   TGAAGAAACCTGGGGGCTTTGACATATCCTTGTTCTACAGAGATATCATCAGCATAGCAGAGGATGAGGA
   CCTCAGGGTTCACTTTGAGGAATCCAGCAAGCTAGAAGACCTGTTGCGGAAGGTTCGCGCCAAGGAGACC
35 AGGAAGCGAGCACTCAGCAGGTTAAAGCTGAAGCTCAACAAAGATATAGTGATCTCTGTGGGCATTTATA
   ATCTGGTCCAGAAGGCTCTCAAGCCTCCTCCAATAAAGCTCTATCGGGAAACAAATGAACCAGTGAAAAC
   CAAGACCCGGACCTTTAATACAAGTACAGGCGGTTTGCTTCTGCCTAGCGATACCAAGAGGTCTCAGATC
   TATGGGAGTCGTCAGATTATACTGGAGAAAGAGGAAACAGAAGAGCTAAAACGGTTTGATGATCCAGGTT
   TGATGCTCATGGGTTTCAAGCCGTTGGTACTGCTGAAGAAACACCATTACCTGAGGCCCTCCCTGTTCGT
40 GTACCCAGAGGAGTCGCTGGTGATTGGGAGCTCAACCCTGTTCAGTGCTCTGCTCATCAAGTGTCTGGAG
   AAGGAGGTTGCAGCATTGTGCAGATACACACCCCGCAGGAACATCCCTCCTTATTTTGTGGCTTTGGTGC
   CACAGGAAGAAGAGTTGGATGACCAGAAAATTCAGGTGACTCCTCCAGGCTTCCAGCTGGTCTTTTTACC
   CTTTGCTGATGATAAAAGGAAGATGCCCTTTACTGAAAAAATCATGGCAACTCCAGAGCAGGTGGGCAAG
   ATGAAGGCTATCGTTGAGAAGCTTCGCTTCACATACAGAAGTGACAGCTTTGAGAACCCCGTGCTGCAGC
45 AGCACTTCAGGAACCTGGAGGCCTTGGCCTTGGATTTGATGGAGCCGGAACAAGCAGTGGACCTGACATT
   GCCCAAGGTTGAAGCAATGAATAAAAGACTGGGCTCCTTGGTGGATGAGTTTAAGGAGCTTGTTTACCCA
   CCAGATTACAATCCTGAAGGGAAAGTTACCAAGAGAAAACACGATAATGAAGGTTCTGGAAGCAAAAGGC
   CCAAGGTGGAGTATTCAGAAGAGGAGCTGAAGACCCACATCAGCAAGGGTACGCTGGGCAAGTTCACTGT
   GCCCATGCTGAAAGAGGCCTGCCGAGCTTACGGGCTGAAGAGTGGGCTGAAGAAGCAGGAGCTGCTGGAA
50 GCCCTCACCAAGCACTTCCAGGACTGACCAGAGGCCGCGCGTCCAGCTGCCCTTCCGCAGTGTGCCAGGC
   TGCCTGGCCTTGTCCTCAGCCAGTTAAAATGTGTTTCTCCTGAGCTAGGAAGAGTCTACCCGACATAAGT
   CGAGGGACTTTATGTTTTTGAGGCTTTCTGTTGCCATGGTGATGGTGTAGCCCTCCCACTTTGCTGTTCC
   TTACTTTACTGCCTGAATAAAGAGCCCTAAGTTTGTACT
```

>gi|4503841|ref|NP_001460.1| thyroid autoantigen 70kDa (Ku antigen);
thyroid autoantigen 70kD (Ku antigen) [Homo sapiens] (SEQ ID NO:68)
```
MSGWESYYKTEGDEEAEEEQEENLEASGDYKYSGRDSLIFLVDASKAMFESQSEDELTPFDMSIQCIQSV
YISKIISSDRDLLAVVFYGTEKDKNSVNFKNIYVLQELDNPGAKRILELDQFKGQQGQKRFQDMMGHGSD
YSLSEVLWVCANLFSDVQFKMSHKRIMLFTNEDNPHGNDSAKASRARTKAGDLRDTGIFLDLMHLKKPGG
FDISLFYRDIISIAEDEDLRVHFEESSKLEDLLRKVRAKETRKRALSRLKLKLNKDIVISVGIYNLVQKA
LKPPPIKLYRETNEPVKTKTRTFNTSTGGLLLPSDTKRSQIYGSRQIILEKEETEELKRFDDPGLMLMGF
KPLVLLKKHHYLRPSLFVYPEESLVIGSSTLFSALLIKCLEKEVAALCRYTPRRNIPPYFVALVPQEEEL
```

DDQKIQVTPPGFQLVFLPFADDKRKMPFTEKIMATPEQVGKMKAIVEKLRFTYRSDSFENPVLQQHFRNL
EALALDLMEPEQAVDLTLPKVEAMNKRLGSLVDEFKELVYPPDYNPEGKVTKRKHDNEGSGSKRPKVEYS
EEELKTHISKGTLGKFTVPMLKEACRAYGLKSGLKKQELLEALTKHFQD

5   >gi|20070134|ref|NM_001469.2| Homo sapiens thyroid autoantigen 70kDa (Ku
    antigen) (G22P1), mRNA (SEQ ID NO:69)
    GCGGGCCGTTATCCATTTGTGTTGTTCGCCAGCTAGGCCTGGCCTCGTCCCGCTTCGCTCGGTCGGTCTC
    GCGCGCCCCCATAGCCTTGCTAGAGGGTTAGCGTTAGCCTTAAGTGTGCGAATCCGAGGAGCAGCGACAG
    ACTCGAGACCACGCTCCTTCCTCGGGAAGGAGGCGGCACCTCGCGTTTGAGGCCCGCCTGCGTTTGAGGC
10 CCGCCTGCGCTTGCGGCCCGCCTGCGCTTGAGGCCTGTCTGCGTTTGAGATCTCATTGGGCGTGATTGAG
    GAATTTGGGGAGGTTTTTGGGCGGTATTGAGGACGAGGGGGTCCGTTAGTCAGCATAGAATCCTGGAGCG
    GGAATCCCTCACCGTCTAAATGGCGTCGGGGGCGGGACCTCCGGGATCTGGCTTCCGCGGGCCGCCGCCG
    GCCCTGAAACGTGAGGGATAGCTGAGATGAGGCAGCTACTGGGATGGCCCCCATGCGCATTTACATGCAG
    TCCGACTGCCGAGCTTTCGAGGCAGCAGGATTTACCGTCCACATTCCTCACTACTAACCAAGCTTTTAGA
15 ACAGATCTCACAAGAACCTAGAGGTCGGTATTTTTTCGATTTAAATTTGCCTGTTACTGACGTTAACGTC
    TTTCGCCTAGTGAGCAGTAGCCAAC<u>ATG</u>TCAGGGTGGGAGTCATATTACAAAACCGAGGGCGATGAAGAA
    GCAGAGGAAGAACAAGAAGAGAACCTTGAAGCAAGTGGAGACTATAAATATTCAGGAAGAGATAGTTTGA
    TTTTTTTGGTTGATGCCTCCAAGGCTATGTTTGAATCTCAGAGTGAAGATGAGTTGACACCTTTTGACAT
    GAGCATCCAGTGTATCCAAAGTGTGTACATCAGTAAGATCATAAGCAGTGATCGAGATCTCTTGGCTGTG
20 GTGTTCTATGGTACCGAGAAAGACAAAAATTCAGTGAATTTTAAAAATATTTACGTCTTACAGGAGCTGG
    ATAATCCAGGTGCAAAACGAATTCTAGAGCTTGACCAGTTTAAGGGGCAGCAGGGACAAAAACGTTTCCA
    AGACATGATGGGCCACGGATCTGACTACTCACTCAGTGAAGTGCTGTGGGTCTGTGCCAACCTCTTTAGT
    GATGTCCAATTCAAGATGAGTCATAAGAGGATCATGCTGTTCACCAATGAAGACAACCCCCATGGCAATG
    ACAGTGCCAAAGCCAGCCGGGCCAGGACCAAAGCCGGTGATCTCCGAGATACAGGCATCTTCCTTGACTT
25 GATGCACCTGAAGAAACCTGGGGGCTTTGACATATCCTTGTTCTACAGAGATATCATCAGCATAGCAGAG
    GATGAGGACCTCAGGGTTCACTTTGAGGAATCCAGCAAGCTAGAAGACCTGTTGCGGAAGGTTCGCGCCA
    AGGAGACCAGGAAGCGAGCACTCAGCAGGTTAAAGCTGAAGCTCAACAAAGATATAGTGATCTCTGTGGG
    CATTTATAATCTGGTCCAGAAGGCTCTCAAGCCTCCTCCAATAAAGCTCTATCGGGAAACAAATGAACCA
    GTGAAAACCAAGACCCGGACCTTTAATACAAGTACAGGCGGTTTGCTTCTGCCTAGCGATACCAAGAGGT
30 CTCAGATCTATGGGAGTCGTCAGATTATACTGGAGAAAGAGGAAACAGAAGAGCTAAAACGGTTTGATGA
    TCCAGGTTTGATGCTCATGGGTTTCAAGCCGTTGGTACTGCTGAAGAAACACCATTACCTGAGGCCCTCC
    CTGTTCGTGTACCCAGAGGAGTCGCTGGTGATTGGGAGCTCAACCCTGTTCAGTGCTCTGCTCATCAAGT
    GTCTGGAGAAGGAGGTTGCAGCATTGTGCAGATACACACCCCGCAGGAACATCCCTCCTTATTTTGTGGC
    TTTGGTGCCACAGGAAGAAGAGTTGGATGACCAGAAAATTCAGGTGACTCCTCCAGGCTTCCAGCTGGTC
35 TTTTTACCCTTTGCTGATGATAAAAGGAAGATGCCCTTTACTGAAAAAATCATGGCAACTCCAGAGCAGG
    TGGGCAAGATGAAGGCTATCGTTGAGAAGCTTCGCTTCACATACAGAAGTGACAGCTTTGAGAACCCCGT
    GCTGCAGCAGCACTTCAGGAACCTGGAGGCCTTGGCCTTGGATTTGATGGAGCCGGAACAAGCAGTGGAC
    CTGACATTGCCCAAGGTTGAAGCAATGAATAAAAGACTGGGCTCCTTGGTGGATGAGTTTAAGGAGCTTG
    TTTACCCACCAGATTACAATCCTGAAGGGAAAGTTACCAAGAGAAAACGATAATGAAGGTTCTGGAAG
40 CAAAAGGCCCAAGGTGGAGTATTCAGAAGAGGAGCTGAAGACCCACATCAGCAAGGGTACGCTGGGCAAG
    TTCACTGTGCCCATGCTGAAAGAGGCCTGCCGGGCTTACGGGCTGAAGAGTGGTCTGAAGAAGCAGGAGC
    TGCTGGAAGCCCTCACCAAGCACTTCCAGGACT<u>GA</u>CCAGAGGCCGCGCGTCCAGCTGCCCTTCCGCAGTG
    TGGCCAGGCTGCCTGGCCTTGTCCTCAGCCAGTTAAAATGTGTTTCTCCTGAGCTAGGAAGAGTCTACCC
    GACATAAGTCGAGGGACTTTATGTTTTTGAGGCTTTCTGTTGCCATGGTGATGGTGTAGCCCTCCCACTT
45 TGCTGTTCTTTACTTTACTGCCTGAATAAAGAGCCCTAAGTTTGTACTAAAAAAAAAAAAAAAAAAAAAA
    AAAAAAAAAAAA

    SCAMP2
50
    >gi|5230677|gb|AF005038.2|AF005038 Homo sapiens secretory carrier membrane
    protein (SCAMP2) mRNA, complete cds (SEQ ID NO:70)
    GCGGAGTTCGCCGCTGGCCCCCGATCACC<u>ATG</u>TCGGCTTTCGACACCAACCCCTTCGCGGACCCAGTGGA
    TGTAAACCCCTTCCAGGATCCCTCTGTGACCCAGCTGACCAACGCCCCGCAGGGCGGCCTGGCGGAATTC
55 AACCCCTTCTCAGAGACAAATGCAGCGACAACAGTTCCTGTCACCCAACTCCCTGGGTCCTCACAGCCAG
    CGGTTCTCCAGCCATCAGTGGAACCAACCCAGCCGACCCCCCAGGCCGTGGTGTCTGCAGCCCAGGCAGG
    CCTGCTCCGGCAGCAGGAAGAACTGGACAGGAAAGCTGCCGAGCTGGAACGCAAGGAGCGGGAGCTGCAG
    AACACTGTAGCCAACTTGCATGTGAGACAGAACAACTGGCCCCCTCTGCCCTCGTGGTGCCCTGTGAAGC
    CCTGCTTCTATCAGGATTTCTCCACAGAGATCCCTGCCGACTACCAGCGGATATGCAAGATGCTCTACTA
60 TCTGTGGATGTTGCATTCAGTGACTCTGTTTCTGAACCTGCTTGCCTGCCTGGCCTGGTTCTCGGGCAAC
    AGCTCCAAGGGAGTGGACTTTGGCCTCTCCATCCTGTGGTTTCTGATCTTCACTCCCTGTGCCTTCCTTT

```
GTTGGTACCGACCCATCTATAAGGCCTTTAGGTCCGACAACTCTTTCAGCTTCTTTGTGTTCTTCTTTGT
ATTTTTTTGTCAAATAGGGATCTACATCATCCAGTTGGTTGGCATCCCTGGCCTGGGGGACAGCGGTTGG
ATTGCAGCCCTGTCTACACTGGATAATCATTCCCTGGCCATATCAGTCATCATGATGGTTGTGGCTGGCT
TCTTCACCCTCTGTGCCGTGCTCTCAGTCTTCCTCCTGCAGCGGGTGCACTCCCTCTACCGACGGACAGG
```
5
```
GGCCAGCTTCCAGCAGGCCCAGGAGGAGTTTTCCCAGGGCATCTTCAGCAGCAGAACCTTCCACAGAGCT
GCTTCATCTGCTGCCCAAGGAGCCTTCCAGGGGAATTAGTCCTCCTCTCTTCTCTCCCCCTCAGCCTTTC
TCTCGCCTGCCTTCTGAGCTGCACTTTCCGTGGGTGCCTTATGTGGTGGTGGTTGTGCCCAGCACAGACC
TGGCAGGGTTCTTGCCGTGGCTCTTCCTCCTCCCTCAGCGACCAGCTCTCCCTGGAACGGGAGGGACAGG
GAATTTTTTCCCCCTCTATGTACAAAAAAAAACAAAGCTCTCTTTCCTTCTCTGGT
```
10

>gi|5230678|gb|AAB62723.2| secretory carrier membrane protein [Homo sapiens] (SEQ ID NO:71)

MSAFDTNPFADPVDVNPFQDPSVTQLTNAPQGGLAEFNPFSETNAATTVPVTQLPGSSQPAVLQPSVEPT
QPTPQAVVSAAQAGLLRQQEELDRKAAELERKERELQNTVANLHVRQNNWPPLPSWCPVKPCFYQDFSTE

15
IPADYQRICKMLYYLWMLHSVTLFLNLLACLAWFSGNSSKGVDFGLSILWFLIFTPCAFLCWYRPIYKAF
RSDNSFSFFVFFFVFFCQIGIYIIQLVGIPGLGDSGWIAALSTLDNHSLAISVIMMVVAGFFTLCAVLSV
FLLQRVHSLYRRTGASFQQAQEEFSQGIFSSRTFHRAASSAAQGAFQGN

20   FIBULIN-5

>gi|5453649|ref|NM_006329.1| Homo sapiens fibulin 5 (FBLN5), mRNA (SEQ ID NO:72)

```
ACCCCGGCGCTCTCCCCGTGTNCTCTCCACGACTCGCTCGGCCCCTCTGGAATAAAACACCCGCGAGCCC
```
25
```
CGAGGGCCCAGAGGAGGCCGACGTGCCCGAGCTCCTCCGGGGGTCCCGCCCGCAAGCTTTCTTCTCGCCT
TCGCATCTCCTCCTCGCGCGTCTTGGACATGCCAGGAATAAAAAGGATACTCACTGTTACCATTCTGGCT
CTCTGTCTTCCAAGCCCTGGGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAG
GACAGTGTTTAGATATTGATGAATGCCGAACCATCCCCGAGGCCTGCCGAGGAGACATGATGTGTGTTAA
```
30
```
CCAAAATGGGGGGTATTTATGCCATTCCCGGACAAACCCTGTGTATCGAGGGCCCTACTCGAACCCCTAC
TCGACCCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAAACTATCCCACGATCT
CCAGGCCTCTTATATGCCGCTTTGGATACCAGATGGATGAAAGCAACCAATGTGTGGATGTGGACGAGTG
TGCAACAGATTCCCACCAGTGCAACCCCACCCAGATTTGCATCAATATGAAGGGCGGGTACACCTGCTCC
```
35
```
TGCACCGACGGATATTGGCTTTTGGAAGGCCAGTGCTTAGACATTGATGAATGTCGCTATGGTTACTGCC
AGCAGCTCTGTGCGAATGTTCCTGGATCCTATTCTTGTACATGCAACCCTGGTTTTACCCTCAATGAGGA
TGGAAGGTCTTGCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAAACCTGCGTCAACACC
TACGGCTCTTTCATCTGCCGCTGTGACCCAGGATATGAACTTGAGGAAGATGGCGTTCATTGCAGTGATA
TGGACGAGTGCAGCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGTGAACCAGCCCGGCACATACTTCTG
```
40
```
CTCCTGCCCTCCAGGCTACATCCTGCTGGATGACAACCGAAGCTGCCAAGACATCAACGAATGTGAGCAC
AGGAACCACACGTGCAACCTGCAGCAGACGTGCTACAATTTACAAGGGGGCTTCAAATGCATCGACCCCA
TCCGCTGTGAGGAGCCTTATCTGAGGATCAGTGATAACCGCTGTATGTGTCCTGCTGAGAACCCTGGCTG
CAGAGACCAGCCCTTTACCATCTTGTACCGGGACATGGACGTGGTGTCAGGACGCTCCGTTCCGCTGAC
ATCTTCCAAATGCAAGCCACGACCCGCTACCCTGGGGCCTATTACATTTTCCAGATCAAATCTGGGAATG
```
45
```
AGGGCAGAGAATTTTACATGCGGCAAACGGGCCCCATCAGTGCCACCCTGGTGATGACACGCCCCATCAA
AGGGCCCCGGGAAATCCAGCTGGACTTGGAAATGATCACTGTCAACACTGTCATCAACTTCAGAGGCAGC
TCCGTGATCCGACTGCGGATATATGTGTCGCAGTACCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCT
GCCTCTCATTGGCACCAAGGGACAGGAGAAGAGAGGAAATAACAGAGAGAATGAGAGCGACACAGACGTT
AGGCATTTCCTGCTGAACGTTTCCCCGAAGAGTCAGCCCCGACTTCCTGACTCTCACCTGTACTATTGCA
```
50
```
GACCTGTCACCCTGCAGGACTTGCCACCCCCAGTTCCTATGACACAGTTATCAAAAAGTATTATCATTGC
TCCCCTGATAGAAGATTGTTGGTGAATTTTCAAGGCCTTCAGTTTATTTCCACTATTTTCAAAGAAAATA
GATTAGGTTTGCGGGGGTCTGAGTCTATGTTCAAAGACTGTGAACAGCTTGCTGTCACTTCTTCACCTCT
TCCACTCCTTCTCTCACTGTGTTACTGCTTTGCAAAGACCCGGGGAGCTGGCGGGGAAACCCTGGGGAGT
AGCTAGTTTGCTTTTTGCGTACACAGAAGAAGGCTATGTAAACAAACCACAGCAGGATC
```


>gi|5453650|ref|NP_006320.1| fibulin 5; urine p50 protein; developmental
55   arteries and neural crest epidermal growth factor-like [Homo sapiens] (SEQ
ID NO:73)

MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQNGGYLCHS
RTNPVYRGPYSNPYSTPYSGPYPAAAPPLSAPNYPTISRPLICRFGYQMDESNQCVDVDECATDSHQCNP
TQICINMKGGYTCSCTDGYWLLEGQCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEDGRSCQDVNE
60   CATENPCVQTCVNTYGSFICRCDPGYELEEDGVHCSDMDECSFSEFLCQHECVNQPGTYFCSCPPGYILL
DDNRSCQDINECEHRNHTCNLQQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPFTILY

RDMDVVSGRSVPADIFQMQATTRYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPREIQLDL
EMITVNTVINFRGSSVIRLRIYVSQYPF

>gi|19743802|ref|NM_006329.2| Homo sapiens fibulin 5 (FBLN5), mRNA (SEQ ID
NO:74)
TTCTGCCCGGGCGCTCGCAGCCGAGCGCGGCCGGGGAAGGGCTCTCCTCCCAGCGCCGAGCACTGGGCCC
TGGCAGACGCCCCAAGATTGTTGTGAGGAGTCTAGCCAGTTGGTGAGCGCTGTAATCTGAACCAGCTGTG
TCCAGACTGAGGCCCCATTTGCATTGTTTAACATACTTAGAAAATGAAGTGTTCATTTTTAACATTCCTC
CTCCAATTGGTTTAATGCTGAATTACTGAAGAGGGCTAAGCAAAACCAGGTGCTTGCGCTGAGGGCTCTG
CAGTGGCTGGGAGGACCCCGGCGCTCTCCCCGTGTCCTCTCCACGACTCGCTCGGCCCCTCTGGAATAAA
ACACCCGCGAGCCCCGAGGGCCCAGAGGAGGCCGACGTGCCCGAGCTCCTCCGGGGGTCCCGCCCGCGAG
CTTTCTTCTCGCCTTCGCATCTCCTCCTCGCGCGTCTTGGAC<u>ATG</u>CCAGGAATAAAAAGGATACTCACTG
TTACCATTCTGGCTCTCTGTCTTCCAAGCCCTGGGAATGCACAGGCACAGTGCACGAATGGCTTTGACCT
GGATCGCCAGTCAGGACAGTGTTTAGATATTGATGAATGCCGAACCATCCCCGAGGCCTGCCGAGGAGAC
ATGATGTGTGTTAACCAAAATGGCGGGTATTTATGCATTCCCCGGACAAACCCTGTGTATCGAGGGCCCT
ACTCGAACCCCTACTCGACCCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAAA
CTATCCCACGATCTCCAGGCCTCTTATATGCCGCTTTGGATACCAGATGGATGAAAGCAACCAATGTGTG
GATGTGGACGAGTGTGCAACAGATTCCACCAGTGCAACCCCACCCAGATCTGCATCAATACTGAAGGCG
GGTACACCTGCTCCTGCACCGACGGATATTGGCTTCTGGAAGGCCAGTGCTTAGACATTGATGAATGTCG
CTATGGTTACTGCCAGCAGCTCTGTGCGAATGTTCCTGGATCCTATTCTTGTACATGCAACCCTGGTTTT
ACCCTCAATGAGGATGGAAGGTCTTGCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAAA
CCTGCGTCAACACCTACGGCTCTTTCATCTGCCGCTGTGACCCAGGATATGAACTTGAGGAAGATGGCGT
TCATTGCAGTGATATGGACGAGTGCAGCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGTGAACCAGCCC
GGCACATACTTCTGCTCCTGCCCTCCAGGCTACATCCTGCTGGATGACAACCGAAGCTGCCAAGACATCA
ACGAATGTGAGCACAGGAACCACACGTGCAACCTGCAGCAGACGTGCTACAATTTACAAGGGGGCTTCAA
ATGCATCGACCCCATCCGCTGTGAGGAGCCTTATCTGAGGATCAGTGATAACCGCTGTATGTGTCCTGCT
GAGAACCCTGGCTGCAGAGACCAGCCCTTTACCATCTTGTACCGGGACATGGACGTGGTGTCAGGACGCT
CCGTTCCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCCTGGGGCCTATTACATTTTCCAGAT
CAAATCTGGGAATGAGGGCAGAGAATTTTACATGCGGCAAACGGGCCCCATCAGTGCCACCCTGGTGATG
ACACGCCCCATCAAAGGGCCCCGGGAAATCCAGCTGGACTTGGAAATGATCACTGTCAACACTGTCATCA
ACTTCAGAGGCAGCTCCGTGATCCGACTGCGGATATATGTGTCGCAGTACCCATTC<u>TGA</u>GCCTCGGGCTG
GAGCCTCCGACGCTGCCTCTCATTGGCACCAAGGGACAGGAGAAGAGAGGAAATAA<u>C</u>AGAGAGAATGAGA
GCGACACAGACGTTAGGCATTTCCTGCTGAACGTTTCCCCGAAGAGTCAGCCCCGACTTCCTGACTCTCA
CCTGTACTATTGCAGACCTGTCACCCTGCAGGACTTGCCACCCCCAGTTCCTATGACACAGTTATCAAAA
AGTATTATCATTGCTCCCCTGATAGAAGATTGTTGGTGAATTTTCAAGGCCTTCAGTTTATTTCCACTAT
TTTCAAAGAAAATAGATTAGGTTTGCGGGGGTCTGAGTCTATGTTCAAAGACTGTGAACAGCTTGCTGTC
ACTTCTTCACCTCTTCCACTCCTTCTCTCACTGTGTTACTGCTTTGCAAAGACCCGGGAGCTGGCGGGGA
ACCCTGGGAGTAGCTAGTTTGCTTTTTGCGTACACAGAGAAGGCTATGTAAACAAACCACAGCAGGATCG
AAGGGTTTTTAGAGAATGTGTTTCAAAACCATGCCTGGTATTTTCAACCATAAAAGAAGTTTCAGTTGTC
CTTAAATTTGTATAACGGTTTAATTCTGTCTTGTTCATTTTGAGTATTTTTAAAAAATATGTCGTAGAAT
TCCTTCGAAAGGCCTTCAGACACATGCTATGTTCTGTCTTCCCAAACCCAGTCTCCTCTCCATTTTAGCC
CAGTGTTTTCTTTGAGGACCCCTTAATCTTGCTTTCTTTAGAATTTTTACCCAATTGGATTGGAATGCAG
AGGTCTCCAAACTGATTAAATATTTGAAGAGAAAAAAAAAAAAAAAAAAAAAAAA

>gi|19743803|ref|NP_006320.2| fibulin 5 precursor; urine p50 protein;
developmental arteries and neural crest epidermal growth factor-like [Homo
sapiens] (SEQ ID NO:75)
MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQNGGYLCIP
RTNPVYRGPYSNPYSTPYSGPYPAAAPPLSAPNYPTISRPLICRFGYQMDESNQCVDVDECATDSHQCNP
TQICINTEGGYTCSCTDGYWLLEGQCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEDGRSCQDVNE
CATENPCVQTCVNTYGSFICRCDPGYELEEDGVHCSDMDECSFSEFLCQHECVNQPGTYFCSCPPGYILL
DDNRSCQDINECEHRNHTCNLQQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPFTILY
RDMDVVSGRSVPADIFQMQATTRYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPREIQLDL
EMITVNTVINFRGSSVIRLRIYVSQYPF

KIAA1228

>gi|20521803|dbj|AB033054.2| Homo sapiens mRNA for KIAA1228 protein,
partial cds (SEQ ID NO:76)
<u>G</u>CGCGGGGCTGAGCCGCCCGGGATCAGCGCGAGCACCCAGCCCGCCTCGGCCGGGAGGGCCTGAGAACCC

CGGGGGCGTGCTGAGCGTGGAGCTGCCCGGGCTGCTGGCGCAGCTGGCGCGGAGCTTCGCGCTGCTGCTG
CCCGTGTACGCGCTGGGCTACCTGGGGCTCAGCTTCAGCTGGGTTCTCCTCGCGCTCGCGCTGCTCGCCT
GGTGTCGCCGCAGCCGCGGCCTCAAGGCCCTGCGCCTGTGCCGCGCGCTGGCGCTGCTGGAAGACGAGGA
GCGCGTCGTGCGCCTGGGGGTGCGCGCCTGCGACCTGCCCGCCTGGGTTCATTTTCCAGACACTGAAAGA
GCAGAATGGCTAAATAAGACTGTAAAACACATGTGGCCTTTCATTTGCCAATTTATAGAGAAGTTGTTTC
GAGAAACTATAGAACCAGCCGTGCGGGGAGCAAACACCCACCTTAGCACCTTTAGTTTCACGAAGGTCGA
CGTGGGCCAGCAGCCCCTCAGGATCAATGGTGTTAAGGTATACACTGAAAATGTAGACAAAAGGCAAATT
ATTTTGGACCTTCAGATTAGTTTTGTAGGAAATTGTGAGATTGATTTGGAGATCAAACGATATTTTTGTA
GAGCTGGTGTGAAAAGTATCCAGATTCATGGTACCATGCGGGTGATCCTGGAACCGTTGATTGGAGATAT
GCCCTTAGTTGGAGCTTTGTCTATCTTCTTCCTTAGGAAACCACTTTTAGAAATTAACTGGACAGGACTG
ACGAATCTTCTGGATGTCCCTGGATTGAATGGTTTATCAGATACTATCATTTTGGATATAATATCAAACT
ATCTGGTGCTTCCCAATCGAATCACCGTTCCACTTGTCAGTGAAGTTCAAATAGCTCAGTTGCGGTTTCC
TGTACCAAAGGGTGTTCTAAGGATACATTTTATTGAAGCTCAGGATCTTCAGGGGAAAGACACTTACCTT
AAGGGACTTGTCAAGGGAAAGTCAGACCCCTATGGAATCATTAGAGTTGGCAACCAAATCTTCCAAAGCA
GAGTCATCAAGGAGAACCTCAGTCCAAAGTGGAATGAAGTCTATGAGGCTTTAGTGTATGAACATCCTGG
ACAAGAATTAGAGATTGAGCTCTTTGATGAAGACCCAGACAAGGATGACTTTTTAGGAAGTCTTATGATT
GACCTCATTGAAGTTGAAAAGGAGCGCCTTTTAGATGAATGGTTCACTCTGGACGAGGTTCCCAAGGGGA
AGCTACACTTGAGACTGGAGTGGCTCACGTTAATGCCAAATGCGTCAAACCTCGACAAGGTGCTAACAGA
CATCAAAGCTGACAAAGACCAAGCCAACGATGGTCTTTCCTCTGCATTGCTGATCTTGTACTTGGATTCA
GCAAGGAACCTTCCGTCAGGGAAGAAAATAAGCAGCAACCCAAATCCTGTTGTCCAGATGTCAGTTGGGC
ACAAGGCCCAGGAGAGCAAGATTCGATACAAAACCAATGAACCTGTGTGGGAGGAAAACTTCACTTTCTT
CATTCACAATCCCAAGCGCCAGGACCTTGAAGTTGAGGTCAGAGACGAGCAGCACCAGTGTTCCCTGGGG
AACCTGAAGGTCCCCCTCAGCCAGCTGCTCACCAGTGAGGACATGACTGTGAGCCAGCGCTTCCAGCTCA
GTAACTCGGGTCCAAACAGCACCATCAAGATGAAGATTGCCCTGCGGGTGCTCCATCTCGAAAAGCGAGA
AAGGCCTCCAGACCACCAACACTCAGCTCAAGTCAAACGTCCCTCTGTGTCCAAAGAGGGGAGGAAAACA
TCCATCAAATCTCATATGTCTGGGTCTCCAGGCCCTGGTGGCAGCAACACAGCTCCATCCACACCAGTCA
TTGGGGGCAGTGATAAGCCTGGTATGGAAGAAAAGGCCCAGCCCCCTGAGGCCGGCCCTCAGGGGCTGCA
CGACCTGGGCAGAAGCTCCTCCAGCCTCCTGGCCTCCCCAGGCCACATCTCAGTCAAGGAGCCGACCCCC
AGCATCGCCTCGGACATCTCGCTGCCCATCGCCACCCAGGAGCTGCGGCAAAGGCTGAGGCAGCTGGAAA
ACGGGACGACCCTGGGACAGTCTCCACTGGGGCAGATCCAGCTGACCATCCGGCACAGCTCGCAGAGAAA
CAAGCTTATCGTGGTCGTGCATGCCTGCAGAAACCTCATTGCCTTCTCTGAAGACGGCTCTGACCCCTAT
GTCCGCATGTATTTATTACCAGACAAGAGGCGGTCAGGAAGGAGGAAAACACACGTGTCAAAGAAAACAT
TAAATCCAGTGTTTGATCAAAGCTTTGATTTCAGTGTTTCGTTACCAGAAGTCCAGAGGAGAACGCTTGA
CGTTGCCGTGAAGAACAGTGGCGGCTTCCTGTCCAAAGACAAAGGGCTCCTTGGCAAAGTATTGGTTGCT
CTGGCATCTGAAGAACTTGCCAAAGGCTGGACCCAGTGGTATGACCTCACGGAAGATGGGACGAGGCCTC
AGGCGATGACATAGCCGCAGCAGGCAGGAGGCGTCCTCTTCAGCGTAGCTCTCCACCTCTACCCGGAACA
CACCCTCTCACAGACGTACCAATGTTATTTTTATAATTTCATGGATTTAGTTATACATACCTTAATAGTT
TTATAAAATTGTTGACATTTCAGGCAAATTTGGCCAATATTATCATTGAATTTTCTGTGTTGGATTTCCT
CTAGGATTTCGCCAGTTCCTACAACGTGCAGTAGGGCGGCGGTAGCTCTTGTGTCTGTGGACTCTGCTCA
GCTGTGTCCGTAGGAGTCGGATGTGTCTGTGCTTTATTATGGCCTTGTTTATATATCACTGAGGTATACT
ATGCCATGTAAATAGACTATTTTTTATAATCTTTACATGCTGGTTTAAATTCAGAAGGAAATAGATCAAG
GAAATATATATATTTTCTTCTAAAACTTATTAAATTCGTGTGACAAATAATCATTTTCATCTTGGCAGCA
AAAGTTCTCAGTGACCTATTTTGTGGTGTTTCTTTTTGAAAAGAAAAGCTGAAATATTATTAAATGCTAG
TATGTTTCTGCCCATTATGAAAGATGAAATAAAGTATTCAAAATATTAACATTTTCATAAATATAAGGAT
GTATTATTGAGAAGTAAGTTGAAGGGCTTATAAGGAAAAATGTTTTATAACTGAGTAATATATTAAGAGA
ATTGTCATGGTTCATAAATCACATTATGCTAATCTGAAATTTCTTACATAAAAATGAAGTGTCTTATGTT
TATTTTAATTGCTGTTGTAACTTACTCATGAAACAGTATACAGACACCTTGTACTTTTCCTCAACTGTAA
GAGCAGACTTTCAATGTTAGCAATTAAGCTGTTGTCAAACAATCAGTCATGAGCTTTGTTAATTTTCAAT
GTTTTCCCAGCCTATAAAAAAGGAAGGTACACGTTGTCCTTTTAAAGGTTGTGAGGTAGATTGGAGTGA
GTAGACAGGATATTGCATTAAAAATTGAAAGCTCGATCTCATTATTGTCAGGAACCCCCAGTGTGACCTC
ACACATAGGATGTGGGACCTTTGAGCCGATGTGCACTGGCCACCACCAGGGTTGGGGGCGCCACAGCTGC
GAGCCCGGCCCCTGCTGTTCTCAGCAGCCACTTCCCAGGCTGCCTCACTTTATGCCATGACTGACCTTAA
TATTGGGATATTGTTATGCAATTTTTATCCTGTTTAGACTGTTAAAAGCAGGTTTCTGTACTTAAGAGTC
CCCCAGACCTCCTGTGAGGTGAGGCTCTGTTGCAGTGTCGTAGGCTGTGTGTGTCTGTAAAGAAAGAATG
CACATATGTAGACGATTAAGTGTATATTATAAGCTATATGCTGAAAATGGCTTCCATAGCCATGAGAACA
TCTTAAAACTATGTGTAAATATATTAAGGAAAGTATAGCTTTGTAATTTAAATTGGAGCTTTTAGCTTGT
TTCATGGAACTATACAACTTGTGGGTAACTCACATGACCAAACAAACCAAAGTGCCTGTGACGGGGCGGG
TGGCGTCTACCCACCCTCCCTTCTAGCAGATTCTTATTTTGTTTGAATTTATAAACAAGGCTGGTGGCTG
TCTACCCACCCTCCCTTCTAGCAGATTTTGATTTTGTTTCAATTTATAACTTACACTTTGAACCATGGGT
TTACTTATAATGGAGTCTGTAGCTTCACAGCATATTTCATGTAATCATAAAGACCAGTATATTCCCTCCT
GCTGAATGACATGCGACTGTAAAGCCTCTTTATAAACCATTTCCAATGTTAGTATATAGGATTATTTGGG

```
     AAGCGTATCAATACCTTTATAGACAAATACGAACATGTATGCACACAAAACATTTAACTATGGTATTTAT
     GGAAGACAGGTAACAACATTAAATCTAGTTGCTTTCCCTTAGTATTAGATTTGTTGAGGGTTTTTTAAAA
     ATCAGGTCTGTTGAAAGTCTTCTGTCATAATCTATAAAGCAGCAGCACTCATGGAAATTGTAGCATGCCA
     GTAATTTTTACCAACATCCCATACATCTGAGTTCTGCAGTCCAGTGTGTAATCCGCTCCATGTGTATTTT
5    GCTTAATGGAATGCTTTATTTAAGCACTTAGGCAGAGTAGACACAATTAAAGGTACAAAGCCCAGAGGAA
     GTGGTAGAGCAGCACCGTGCCTGCCCTGAGGCAGTGGAGTCAGTAGCGCTGTCCCCAGGGCCTTGAGTGC
     CTGGAGGTGCTTGGCCTCCAGTAGCTGCCTCCATTCTCTTTTAAAAAAGGGGGTGATTCTGAGGCACTG
     AAGTGCCTCCCAGATGTGGAGGAGTGAAGCCACCATCGAGGCCACACTCAGCACTCCAGGATCCCAGCGA
     TGTCAGACACTCTTGAGTTGTCAAAACGTTAATTTTCAGTTTTAAATAATCAGTTTATCTAAGAAAGGG
10   AATTTTAACTTTTCTACCTTGAGCCAAGCCAATGAAGGGAAAATTAATTAACTTAGTAAATTTGAAGTGC
     AGCTCTGTTAGCTCGTACATGTGGGTTCTTATCCTGATCCTGTGCCTTAAAGTAGGAAGGTGTTTCCAAG
     TTCAGATTAAAATAGAAGCAGCTGGCCGGGTGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGC
     CGAGGCGGGCGGATGACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGAGAAACCCCATCTTT
     ACTAAAAATACAAAAATTAGCCGGGCGTGGTGGTGGAGCGCACGCCTGTAGTCCCAGCTACTCGGGAGGC
15   TGAGGCAGGAGAATTGCTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCCAAGATCGCACCACTGCACTCC
     AGCCTGGGCAACAGAGTGAGACTCCGTCTCAAAAAATAAATAAAATAGAAGCAGCCTTGTAACTGTATTT
     ACCATGATAATATATTCTGCACGGTAAGAATTCCTTTTACAGACATTCTTTATCAAGAGGTCGGCCCTTC
     TTTTTCAGGCACATAAGCCAAATGCAGGCCTGTGTGTAGCTGTGTGTTTTTTCTGTGGTTGCCGCATTTA
     TTCCACCTCCAGCTGGACCCCCCACTGCAAATAGAGAACAGCGGTGGGGGATGGGGGTTAAAAAGTAGAG
20   AACCTCCTTTCTGTTCAACTAATTTCACGTGACAGTGCATGTATTTATTCAATAAAACCTTTATGTTAGC
     TC
```

```
>gi|20521804|dbj|BAA86542.2| KIAA1228 protein [Homo sapiens] (SEQ ID NO:77)
AGLSRPGSARAPSPPRPGGPENPGGVLSVELPGLLAQLARSFALLLPVYALGYLGLSFSWVLLALALLAW
CRRSRGLKALRLCRALALLEDEERVVRLGVRACDLPAWVHFPDTERAEWLNKTVKHMWPFICQFIEKLFR
ETIEPAVRGANTHLSTFSFTKVDVGQQPLRINGVKVYTENVDKRQIILDLQISFVGNCEIDLEIKRYFCR
AGVKSIQIHGTMRVILEPLIGDMPLVGALSIFFLRKPLLEINWTGLTNLLDVPGLNGLSDTIILDIISNY
LVLPNRITVPLVSEVQIAQLRFPVPKGVLRIHFIEAQDLQGKDTYLKGLVKGKSDPYGIIRVGNQIFQSR
VIKENLSPKWNEVYEALVYEHPGQELEIELFDEDPDKDDFLGSLMIDLIEVEKERLLDEWFTLDEVPKGK
LHLRLEWLTLMPNASNLDKVLTDIKADKDQANDGLSSALLILYLDSARNLPSGKKISSNPNPVVQMSVGH
KAQESKIRYKTNEPVWEENFTFFIHNPKRQDLEVEVRDEQHQCSLGNLKVPLSQLLTSEDMTVSQRFQLS
NSGPNSTIKMKIALRVLHLEKRERPPDHQHSAQVKRPSVSKEGRKTSIKSHMSGSPGPGGSNTAPSTPVI
GGSDKPGMEEKAQPPEAGPQGLHDLGRSSSSLLASPGHISVKEPTPSIASDISLPIATQELRQRLRQLEN
GTTLGQSPLGQIQLTIRHSSQRNKLIVVVHACRNLIAFSEDGSDPYVRMYLLPDKRRSGRRKTHVSKKTL
NPVFDQSFDFSVSLPEVQRRTLDVAVKNSGGFLSKDKGLLGKVLVALASEELAKGWTQWYDLTEDGTRPQ
AMT
```

```
>gi|22045983|ref|XM_036408.3| Homo sapiens KIAA1228 protein (KIAA1228),
mRNA (SEQ ID NO:78)
GCGCGGGGCTGAGCCGCCCGGGATCAGCGCGAGCACCCAGCCCGCCTCGGCCGGGAGGGCCTGAGAACCC
CGGGGGCGTGCTGAGCGTGGAGCTGCCCGGGCTGCTGGCGCAGCTGGCGCGGAGCTTCGCGCTGCTGCTG
CCCGTGTACGCGCTGGGCTACCTGGGGCTCAGCTTCAGCTGGGTTCTCCTCGCGCTCGCGCTGCTCGCCT
GGTGTCGCCGCAGCCGCGGCCTCAAGGCCCTGCGCCTGTGCCGCGCGCTGGCGCTGCTGGAAGACGAGGA
GCGCGTCGTGCGCCTGGGGGTGCGCGCCTGCGACCTGCCCGCCTGGGTTCATTTTCCAGACACTGAAAGA
GCAGAATGGCTAAATAAGACTGTAAAACACATGTGGCCTTTCATTTGCCAATTTATAGAAGTTGTTTC
GAGAAACTATAGAACCAGCCGTGCGGGGAGCAAACACCCACCTTAGCACCTTTAGTTTCACGAAGGTCGA
CGTGGGCCAGCAGCCCCTCAGGATCAATGGTGTTAAGGTATACACTGAAAATGTAGACAAAAGGCAAATT
ATTTTGGACCTTCAGATTAGTTTTGTAGGAAATTGTGAGATTGATTTGGAGATCAAACGATATTTTGTA
GAGCTGGTGTGAAAGTATCCAGATTCATGGTACCATGCGGGTGATCCTGGAACCGTTGATTGGAGATAT
GCCCTTAGTTGGAGCTTTGTCTATCTTCTTCCTTAGGAAACCACTTTTAGAAATTAACTGGACAGGACTG
ACGAATCTTCTGGATGTCCCTGGATTGAATGGTTTATCAGATACTATCATTTTGGATATAATATCAAACT
ATCTGGTGCTTCCCAATCGAATCACCGTTCCACTTGTCAGTGAAGTTCAAATAGCTCAGTTGCGGTTTCC
TGTACCAAAGGGTGTTCTAAGGATACATTTTATTGAAGCTCAGGATCTTCAGGGGAAAGACACTTACCTT
AAGGGACTTGTCAAGGGAAAGTCAGACCCCTATGGAATCATTAGAGTTGGCAACCAAATCTTCCAAAGCA
GAGTCATCAAGGAGAACCTCAGTCCAAAGTGGAATGAAGTCTATGAGGCTTTAGTGTATGAACATCCTGG
ACAAGAATTAGAGATTGAGCTCTTTGATGAAGACCCAGACAAGGATGACTTTTTAGGAAGTCTTATGATT
GACCTCATTGAAGTTGAAAAGGAGCGCCTTTTAGATGAATGGTTCACTCTGGACGAGGTTCCCAAGGGGA
AGCTACACTTGAGACTGGAGTGGCTCACGTTAATGCCAAATGCGTCAAACCTCGACAAGGTGCTAACAGA
CATCAAAGCTGACAAAGACCAAGCCAACGATGGTCTTTCCTCTGCATTGCTGATCTTGTACTTGGATTCA
GCAAGGAACCTTCCGTCAGGGAAGAAAATAAGCAGCAACCCAAATCCTGTTGTCCAGATGTCAGTTGGGC
ACAAGGCCCAGGAGAGCAAGATTCGATACAAAACCAATGAACCTGTGTGGGAGGAAAACTTCACTTTCTT
```

100

CATTCACAATCCCAAGCGCCAGGACCTTGAAGTTGAGGTCAGAGACGAGCAGCACCAGTGTTCCCTGGGG
AACCTGAAGGTCCCCCTCAGCCAGCTGCTCACCAGTGAGGACATGACTGTGAGCCAGCGCTTCCAGCTCA
GTAACTCGGGTCCAAACAGCACCATCAAGATGAAGATTGCCCTGCGGGTGCTCCATCTCGAAAAGCGAGA
AAGGCCTCCAGACCACCAACACTCAGCTCAAGTCAAACGTCCCTCTGTGTCCAAAGAGGGGAGGAAAACA
5 TCCATCAAATCTCATATGTCTGGGTCTCCAGGCCCTGGTGGCAGCAACACAGCTCCATCCACACCAGTCA
TTGGGGGCAGTGATAAGCCTGGTATGGAAGAAAAGGCCCAGCCCCCTGAGGCCGGCCCTCAGGGGCTGCA
CGACCTGGGCAGAAGCTCCTCCAGCCTCCTGGCCTCCCCAGGCCACATCTCAGTCAAGGAGCCGACCCCC
AGCATCGCCTCGGACATCTCGCTGCCCATCGCCACCCAGGAGCTGCGGCAAAGGCTGAGGCAGCTGGAAA
ACGGGACGACCCTGGGACAGTCTCCACTGGGGCAGATCCAGCTGACCATCCGGCACAGCTCGCAGAGAAA
10 CAAGCTTATCGTGGTCGTGCATGCCTGCAGAAACCTCATTGCCTTCTCTGAAGACGGCTCTGACCCCTAT
GTCCGCATGTATTTATTACCAGACAAGAGGCGGTCAGGAAGGAGGAAAACACACGTGTCAAAGAAAACAT
TAAATCCAGTGTTTGATCAAAGCTTTGATTTCAGTGTTTCGTTACCAGAAGTGCAGAGGAGAACGCTCGA
CGTTGCCGTGAAGAACAGTGGCGGCTTCCTGTCCAAAGACAAAGGGCTCCTTGGCAAAGTATTGGTTGCT
CTGGCATCTGAAGAACTTGCCAAAGGCTGGACCCAGTGGTATGACCTCACGGAAGATGGGACGAGGCCTC
15 AGGCGATGAC<u>A</u>TAGCCGCAGCAGGCAGGAGGCGTCCTCTTCAGCGTAGCTCTCCACCTCTACCCGGAACA
CACCCTCTCACAGACGTACCAATGTTATTTTTATAATTTCATGGATTTAGTTATACATACCTTAATAGTT
TTATAAAATTGTTGACATTTCAGGCAAATTTGGCCAATATTATCATTGAATTTTCTGTGTTGGATTTCCT
CTAGGATTTCGCCAGTTCCTACAACGTGCAGTAGGGCGGCGGTAGCTCTTGTGTCTGTGGACTCTGCTCA
GCTGTGTCCGTAGGAGTCGGATGTGTCTGTGCTTTATTATGGCCTTGTTTATATATCACTGAGGTATACT
20 ATGCCATGTAAATAGACTATTTTTTATAATCTTTACATGCTGGTTTAAATTCAGAAGGAAATAGATCAAG
GAAATATATATATTTTCTTCTAAAACTTATTAAATTCGTGTGACAAATAATCATTTTCATCTTGGCAGCA
AAAAGTTCTCAGTGACCTATTTTGTGGTGTTTCTTTTTGAAAAGAAAAGCTGAAATATTATTAAATGCTA
GTATGTTTCTGCCCATTATGAAAGATGAAATAAAGTATTCAAAATATTAACATTTTCATAAATATAAGGA
TGTATTATTGAGAAGTAAGTTGAAGGGCTTATAAGGAAAAATGTTTTATAACTGAGTAATATATTAAGAG
25 AATTGTCATGGTTCATAAATCACATTATGCTAATCTGAAATTTCTTACATAAAAATGAAGTGTCTTATGT
TTATTTTAATTGCTGTTGTAACTTACTCATGAAACAGTATACAGACACCTTGTACTTTTCCTCAACTGTA
AGAGCAGACTTTCAATGTTAGCAATTAAGCTGTTGTCAAACAATCAGTCATGAGCTTTGTTAATTTTCAA
TGTTTTCCCAGCCTATAAAAAAAGGAAGGTACACGTTGTCCTTTTAAAGGTTGTGAGGTAGATTGGAGTG
AGTAGACAGGATATTGCATTAAAAATTGAAAGCTCGATCTCATTATTGTCAGGAACCCCCAGTGTGACCT
30 CACACATAGGATGTGGGACCTTTGAGCCGATGTGCACTGGCCACCACCAGGGTTGGGGGCGCCACAGCTG
CGAGCCCGGCCCCTGCTGTTCTCAGCAGCCACTTCCCAGGCTGCCTCACTTTATGCCATGACTGACCTTA
ATATTGGGATATTGTTATGCAATTTTTATCCTGTTTAGACTGTTAAAAGCAGGTTTCTGTACTTAAGAGT
CCCCCAGACCTCCTGTGAGGTGAGGCTCTGTTGCAGTGTCGTAGGCTGTGTGTGTCTGTAAAGAAAGAAT
GCACATATGTAGACGATTAAGTGTATATTATAAGCTATATGCTGAAAATGGCTTCCATAGCCATGAGAAC
35 ATCTTAAAACTATGTGTAAATATATTAAGGAAAGTATAGCTTTGTAATTTAAATTGGAGCTTTTAGCTTG
TTTCATGGAACTATACAACTTGTGGGTAACTCACATGACCAAACAAACCAAAGTGCCTGTGACGGGGCGG
GTGGCGTCTACCCACCCTCCCTTCTAGCAGATTCTTATTTTGTTTGAATTTATAAACAAGGCTGGTGGCT
GTCTACCCACCCTCCCTTCTAGCAGATTTTGATTTTGTTTCAATTTATAACTTACACTTTGAACCATGGG
TTTACTTATAATGGAGTCTGTAGCTTCACAGCATATTTCATGTAATCATAAAGACCAGTATATTCCCTCC
40 TGCTGAATGACATGCGACTGTAAAGCCTCTTTATAAACCATTTCCAATGTTAGTATATAGGATTATTTGG
GAAGCGTATCAATACCTTTATAGACAAATACGAACATGTATGCACACAAAACATTTAACTATGGTATTTA
TGGAAGACAGGTAACAACATTAAATCTAGTTGCTTTCCCTTAGTATTAGATTTGTTGAGGGTTTTTTAAA
AATCAGGTCTGTTGAAAGTCTTCTGTCATAATCTATAAAGCAGCAGCACTCATGGAAATTGTAGCATGCC
AGTAATTTTTACCAACATCCCATACATCTGAGTTCTGCAGTCCAGTGTGTAATCCGCTCCATGTGTATTT
45 TGCTTAATGGAATGCTTTATTTAAGCACTTAGGCAGAGTAGACACAATTAAAGGTACAAAGCCCAGAGGA
AGTGGTAGAGCAGCACCGTGCCTGCCCTGAGGCAGTGGAGTCAGTAGCGCTGTCCCCAGGGCCTTGAGTG
CCTGGAGGTGCTTGGCCTCCAGTAGCTGCCTCCATTCTCTTTTAAAAAAAGGGGGTGATTCTGAGGCACT
GAAGTGCCTCCCAGATGTGGAGGAGTGAAGCCACCATCGAGGCCACACTCAGCACTCCAGGATCCCAGCG
ATGTCAGACACTCTTGAGTTGTCAAAACGTTAATTTTCAGTTTTAAATAATCAGTTTATCTAAGAAAAGG
50 GAATTTTAACTTTTCTACCTTGAGCCAAGCCAATGAAGGGAAAATTAATTAACTTAGTAAATTTGAAGTG
CAGCTCTGTTAGCTCGTACATGTGGGTTCTTATCCTGATCCTGTGCCTTAAAGTAGGAAGGTGTTTCCAA
GTTCAGATTAAAATAGAAGCAGCTGGCCGGGTGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGG
CCGAGGCGGGCGGATGACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGAGAAACCCCATCTT
TACTAAAAATACAAAAATTAGCCGGGCGTGGTGGTGGAGCGCACGCCTGTAGTCCCAGCTACTCGGGAGG
55 CTGAGGCAGGAGAATTGCTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCCAAGATCGCACCACTGCACTC
CAGCCTGGGCAACAGAGTGAGACTCCGTCTCAAAAAATAAATAAAATAGAAGCAGCCTTGTAACTGTATT
TACCATGATAATATATTCTGCACGGTAAGAATTCCTTTTACAGACATTCTTTATCAAGAGGTCGGCCCTT
CTTTTTCAGGCACATAAGCCAAATGCAGGCCTGTGTGTAGCTGTGTGTTTTTTCTGTGGTTGCCGCATTT
ATTCCACCTCCAGCTGGACCCCCCACTGCAAATAGAGAACAGCGGTGGGGGATGGGGGTTAAAAAGTAGA
60 GAACCTCCTTTCTGTTCAACTAATTTCACGTGACAGTGCATGTATTTATTCAATAAAACCTTTATGTTAG
CTC

>gi|22045984|ref|XP_036408.3| similar to RIKEN cDNA 4921504I16 [Homo
sapiens] (SEQ ID NO:79)
MWPFICQFIEKLFRETIEPAVRGANTHLSTFSFTKVDVGQQPLRINGVKVYTENVDKRQIILDLQISFVG
5   NCEIDLEIKRYFCRAGVKSIQIHGTMRVILEPLIGDMPLVGALSIFFLRKPLLEINWTGLTNLLDVPGLN
GLSDTIILDIISNYLVLPNRITVPLVSEVQIAQLRFPVPKGVLRIHFIEAQDLQGKDTYLKGLVKGKSDP
YGIIRVGNQIFQSRVIKENLSPKWNEVYEALVYEHPGQELEIELFDEDPDKDDFLGSLMIDLIEVEKERL
LDEWFTLDEVPKGKLHLRLEWLTLMPNASNLDKVLTDIKADKDQANDGLSSALLILYLDSARNLPSGKKI
SSNPNPVVQMSVGHKAQESKIRYKTNEPVWEENFTFFIHNPKRQDLEVEVRDEQHQCSLGNLKVPLSQLL
10   TSEDMTVSQRFQLSNSGPNSTIKMKIALRVLHLEKRERPPDHQHSAQVKRPSVSKEGRKTSIKSHMSGSP
GPGGSNTAPSTPVIGGSDKPGMEEKAQPPEAGPQGLHDLGRSSSSLLASPGHISVKEPTPSIASDISLPI
ATQELRQRLRQLENGTTLGQSPLGQIQLTIRHSSQRNKLIVVVHACRNLIAFSEDGSDPYVRMYLLPDKR
RSGRRKTHVSKKTLNPVFDQSFDFSVSLPEVQRRTLDVAVKNSGGFLSKDKGLLGKVLVALASEELAKGW
TQWYDLTEDGTRPQAMT
15

EST FROM CLONE 2108068

>gi|8249872|emb|AL357532.1|IROEST111 Homo sapiens EST from clone 2108068,
20   full insert (SEQ ID NO:80)
TCAGCATGAAGCAGTTTGCTGAAGGCTCCACTCTCAAACTGGCTAAGCAGTGTCGAAAGTGGCTGTGCAA
TGACCAGATCGACGCAGGCACTCGGCGCTGGGCAGTGGAGGGCCTGGCTTACCTGACCTTTGATGCCGAC
GTGAAGGAAGAGTTTGTGGAGGATGCGGCTGCTCTGAAAGCTCTGTTCCAGCTCAGCAGGTTGGAGGAGA
GGTCAGTGCTCTTTGCGGTGGCCTCAGCGCTGGTGAACTGCACCAACAGCTATGACTACGAGGAGCCCGA
25   CCCCAAGATGGTGGAGCTGGCCAAGTATGCCAAGCAGCATGTGCCCGAGCAGCACCCCAAGGACAAGCCA
AGCTTCGTGCGGGCTCGGGTGAAGAAGCTGCTGGCAGCGGGTGTGGTGTCGGCCATGGTGTGCATGGTGA
AGACGGAGAGCCCTGTGCTGACCAGTTCCTGCAGAGAGCTGCTCTCCAGGGTCTTCTTGGCTTTAGTGGA
AAGGGCGCACACATCAGCAGCCTCACCAGCTGTGAGCCTGCTATCAGGCCTGCCCCTCCAATAAAAGTGT
GTAGAACTCCCAAAAAAAAAAAAAAAAA
30

vimentin

>gi|4507894|ref|NM_003380.1| Homo sapiens vimentin (VIM), mRNA (SEQ ID
35   NO:81)
GGGCGCGCCAGAGACGCAGCCGCGCTCCCACCACCCACACCCACCGCGCCCTCGTTCGCCTCTTCTCCGG
GAGCCAGTCCGCGCCACCGCCGCCGCCCAGGCCATCGCCACCCTCCGCAGCCATGTCCACCAGGTCCGTG
TCCTCGTCCTCCTACCGCAGGATGTTCGGCGGCCCGGGCACCGCGAGCCGGCCGAGCTCCAGCCGGAGCT
ACGTGACTACGTCCACCCGCACCTACAGCCTGGGCAGCGCGCTGCGCCCCAGCACCAGCCGCAGCCTCTA
40   CGCCTCGTCCCCGGGCGGCGTGTATGCCACGCGCTCCTCTGCCGTGCGCCTGCGGAGCAGCGTGCCCGGG
GTGCGGCTCCTGCAGGACTCGGTGGACTTCTCGCTGGCCGACGCCATCAACACCGAGTTCAAGAACACCC
GCACCAACGAGAAGGTGGAGCTGCAGGAGCTGAATGACCGCTTCGCCAACTACATCGACAAGGTGCGCTT
CCTGGAGCAGCAGAATAAGATCCTGCTGGCCGAGCTCGAGCAGCTCAAGGGCCAAGGCAAGTCGCGCCTA
GGGGACCTCTACGAGGAGGAGATGCGGGAGCTGCGCCGGCAGGTGGACCAGCTAACCAACGACAAAGCCC
45   GCGTCGAGGTGGAGCGCGACAACCTGGCCGAGGACATCATGCGCCTCCGGGAGAAATTGCAGGAGGAGAT
GCTTCAGAGAGAGGAAGCCGAAAACACCCTGCAATCTTTCAGACAGGATGTTGACAATGCGTCTCTGGCA
CGTCTTGACCTTGAACGCAAAGTGGAATCTTTGCAAGAAGAGATTGCCTTTTTGAAGAAACTCCACGAAG
AGGAAATCCAGGAGCTGCAGGCTCAGATTCAGGAACAGCATGTCCAAATCGATGTGGATGTTTCCAAGCC
TGACCTCACGGCTGCCCTGCGTGACGTACGTCAGCAATATGAAAGTGTGGCTGCCAAGAACCTGCAGGAG
50   GCAGAAGAATGGTACAAATCCAAGTTTGCTGACCTCTCTGAGGCTGCCAACCGGAACAATGACGCCCTGC
GCCAGGCAAAGCAGGAGTCCACTGAGTACCGGAGACAGGTGCAGTCCCTCACCTGTGAAGTGGATGCCCT
TAAAGGAACCAATGAGTCCCTGGAACGCCAGATGCGTGAAATGGAAGAGAACTTTGCCGTTGAAGCTGCT
AACTACCAAGACACTATTGGCCGCCTGCAGGATGAGATTCAGAATATGAAGGAGGAAATGGCTCGTCACC
TTCGTGAATACCAAGACCTGCTCAATGTTAAGATGGCCCTTGACATTGAGATTGCCACCTACAGGAAGCT
55   GCTGGAAGGCGAGGAGAGCAGGATTTCTCTGCCTCTTCCAAACTTTTCCTCCCTGAACCTGAGGGAAACT
AATCTGGATTCACTCCCTCTGGTTGATACCCACTCAAAAAGGACATTCCTGATTAAGACGGTTGAAACTA
GAGATGGACAGGTTATCAACGAAACTTCTCAGCATCACGATGACCTTGAATAAAAATTGCACACACTCAG
TGGCAGGCGATATATTACCCAGGCAAGAATAAAAAAGAAATCCCATATCTTAAAGAAACAGCTTTCAAGT
GCCTTTCTGCAGTTTTTCAGGAGCGCAAGATAGATTTGGAATAGGAATAAGCTCTAGTTCTTAACAACCG
60   ACACTCCTACAAGATTTAGAAAAAGTTTACAACATAATCTAGTTTACAGAAAATCTTGTGCTAGAATA
CTTTTTAAAAGGTATTTTGAATACCATTAAAACTGCTTTTTTTTTTCCAGCAAGTATCCAACCAACTTGG

TTCTGCTTCAATAAATCTTTGGAAAACTCCA

>gi|4507895|ref|NP_003371.1| vimentin [Homo sapiens] (SEQ ID NO:82)
MSTRSVSSSSYRRMFGGPGTASRPSSSRSYVTTSTRTYSLGSALRPSTSRSLYASSPGGVYATRSSAVRL
RSSVPGVRLLQDSVDFSLADAINTEFKNTRTNEKVELQELNDRFANYIDKVRFLEQQNKILLAELEQLKG
QGKSRLGDLYEEEMRELRRQVDQLTNDKARVEVERDNLAEDIMRLREKLQEEMLQREEAENTLQSFRQDV
DNASLARLDLERKVESLQEEIAFLKKLHEEEIQELQAQIQEQHVQIDVDVSKPDLTAALRDVRQQYESVA
AKNLQEAEEWYKSKFADLSEAANRNNDALRQAKQESTEYRRQVQSLTCEVDALKGTNESLERQMREMEEN
FAVEAANYQDTIGRLQDEIQNMKEEMARHLREYQDLLNVKMALDIEIATYRKLLEGEESRISLPLPNFSS
LNLRETNLDSLPLVDTHSKRTFLIKTVETRDGQVINETSQHHDDLE

filamin A, alpha

>gi|4503744|ref|NM_001456.1| Homo sapiens filamin A, alpha (actin binding
protein 280) (FLNA), mRNA (SEQ ID NO:83)
GCGATCCGGGCGCCACCCCGCGGTCATCGGTCACCGGTCGCTCTCAGGAACAGCAGCGCAACCTCTGCTC
CCTGCCTCGCCTCCCGCGCGCCTAGGTGCCTGCGACTTTAATTAAAGGGCCGTCCCCTCGCCGAGGCTGC
AGCACCGCCCCCCGGCTTCTCGCGCCTCAAA<u>AT</u>GAGTAGCTCCCACTCTCGGGCGGGCCAGAGCGCAGC
AGGCGCGGCTCCGGGCGGCGGCGTCGACACGCGGGACGCCGAGATGCCGGCCACCGAGAAGGACCTGGCG
GAGGACGCGCCGTGGAAGAAGATCCAGCAGAACACTTTCACGCGCTGGTGCAACGAGCACCTGAAGTGCG
TGAGCAAGCGCATCGCCAACCTGCAGACGGACCTGAGCGACGGGCTGCGGCTTATCGCGCTGTTGGAGGT
GCTCAGCCAGAAGAAGATGCACCGCAAGCACAACCAGCGGCCCACTTTCCGCCAAATGCAGCTTGAGAAC
GTGTCGGTGGCGCTCGAGTTCCTGGACCGCGAGAGCATCAAACTGGTGTCCATCGACAGCAAGGCCATCG
TGGACGGGAACCTGAAGCTGATCCTGGGCCTCATCTGGACCCTGATCCTGCACTACTCCATCTCCATGCC
CATGTGGGACGAGGAGGAGGATGAGGAGGCCAAGAAGCAGACCCCCAAGCAGAGGCTCCTGGGCTGGATC
CAGAACAAGCTGCCGCAGCTGCCCATCACCAACTTCAGCCGGGACTGGCAGAGCGGCCGGGCCCTGGGCG
CCCTGGTGGACAGCTGTGCCCCGGGCCTGTGTCCTGACTGGGACTCTTGGGACGCCAGCAAGCCCGTTAC
CAATGCGCGAGAGGCCATGCAGCAGGCGGATGACTGGCTGGGCATCCCCCAGGTGATCACCCCCGAGGAG
ATTGTGGACCCCAACGTGGACGAGCACTCTGTCATGACCTACCTGTCCCAGTTCCCCAAGGCCAAGCTGA
AGCCAGGGGCTCCCTTGCGCCCCAAACTGAACCCGAAGAAAGCCCGTGCCTACGGGCCAGGCATCGAGCC
CACAGGCAACATGGTGAAGAAGCGGGCAGAGTTCACTGTGGAGACCAGAAGTGCTGGCCAGGGAGAGGTG
CTGGTGTACGTGGAGGACCCGGCCGGACACCAGGAGGAGGCAAAAGTGACCGCCAATAACGACAAGAACC
GCACCTTCTCCGTCTGGTACGTCCCCGAGGTGACGGGGACTCATAAGGTTACTGTGCTCTTTGCTGGCCA
GCACATCGCCAAGAGCCCCTTCGAGGTGTACGTGGATAAGTCACAGGGTGACGCCAGCAAAGTGACAGCC
CAAGGTCCCGGCCTGGAGCCCAGTGGCAACATCGCCAACAAGACCACCTACTTTGAGATCTTTACGGCAG
GAGCTGGCACGGGCGAGGTCGAGGTTGTGATCCAGGACCCCATGGGACAGAAGGGCACGGTAGAGCCTCA
GCTGGAGGCCCGGGGCGACAGCACATACCGCTGCAGCTACCAGCCCACCATGGAGGGCGTCCACACCGTG
CACGTCACGTTTGCCGGCGTGCCCATCCCTCGCAGCCCCTACACTGTCACTGTTGGCCAAGCCTGTAACC
CGAGTGCCTGCCGGGCGGTTGGCCGGGGCCTCCAGCCCAAGGGTGTGCGGGTGAAGGAGACAGCTGACTT
CAAGGTGTACACAAAGGGCGCTGGCAGTGGGGAGCTGAAGGTCACCGTGAAGGGCCCCAAGGGAGAGGAG
CGCGTGAAGCAGAAGGACCTGGGGGATGGCGTGTATGGCTTCGAGTATTACCCCATGGTCCCTGGAACCT
ATATCGTCACCATCACGTGGGGTGGTCAGAACATCGGGCGCAGTCCCTTCGAAGTGAAGGTGGGCACCGA
GTGTGGCAATCAGAAGGTACGGGCCTGGGGCCCTGGGCTGGAGGGCGGCGTCGTTGGCAAGTCAGCAGAC
TTTGTGGTGGAGGCTATCGGGGACGACGTGGGCACGCTGGGCTTCTCGGTGGAAGGGCCATCGCAGGCTA
AGATCGAATGTGACGACAAGGGCGACGGCTCCTGTGATGTGCGCTACTGGCCGCAGGAGGCTGGCGAGTA
TGCCGTTCACGTGCTGTGCAACAGCGAAGACATCCGCCTCAGCCCCTTCATGGCTGACATCCGTGACGCG
CCCCAGGACTTCCACCCAGACAGGGTGAAGGCACGTGGGCCTGGATTGGAGAAGACAGGTGTGGCCGTCA
ACAAGCCAGCAGAGTTCACAGTGGATGCCAAGCACGGTGGCAAGGCCCCACTTCGGGTCCAAGTCCAGGA
CAATGAAGGCTGCCCTGTGGAGGCGTTGGTCAAGGACAACGGCAATGGCACTTACAGCTGCTCCTACGTG
CCCAGGAAGCCGGTGAAGCACACAGCCATGGTGTCCTGGGGAGGCGTCAGCATCCCCAACAGCCCCTTCA
GGGTGAATGTGGGAGCTGGCAGCCACCCCAACAAGGTCAAAGTATACGGCCCCGGAGTAGCCAAGACAGG
GCTCAAGGCCCACGAGCCCACCTACTTCACTGTGGACTGCGCCGAGGCTGGCCAGGGGGACGTCAGCATC
GGCATCAAGTGTGCCCCTGGAGTGGTAGGCCCCGCCGAAGCTGACATCGACTTCGACATCATCCGCAATG
ACAATGACACCTTCACGGTCAAGTACACGCCCCGGGGGGCTGGCAGCTACACCATTATGGTCCTCTTTGC
TGACCAGGCCACGCCCACCAGCCCCATCCGAGTCAAGGTGGAGCCCTCTCATGACGCCAGTAAGGTGAAG
GCCGAGGGCCCTGGCCTCAGTCGCACTGGTGTCGAGCTTGGCAAGCCCACCCACTTCACAGTAAATGCCA
AAGCTGCTGGCAAAGGCAAGCTGGACGTCCAGTTCTCAGGACTCACCAAGGGGGATGCAGTGCGAGATGT
GGACATCATCGACCACCATGACAACACCTACACAGTCAAGTACACGCCTGTCCAGCAGGGTCCAGTAGGC
GTCAATGTCACTTATGGAGGGGATCCCATCCCTAAGAGCCCTTTCTCAGTGGCAGTATCTCCAAGCCTGG
ACCTCAGCAAGATCAAGGTGTCTGGCCTGGGAGAGAAGGTGGACGTTGGCAAAGACCAGGAGTTCACAGT

```
CAAATCAAAGGGTGCTGGTGGTCAAGGCAAAGTGGCATCCAAGATTGTGGGCCCCTCGGGTGCAGCGGTG
CCCTGCAAGGTGGAGCCAGGCCTGGGGGCTGACAACAGTGTGGTGCGCTTCCTGCCCCGTGAGGAAGGGC
CCTATGAGGTGGAGGTGACCTATGACGGCGTGCCCGTGCCTGGCAGCCCCTTTCCTCTGGAAGCTGTGGC
CCCCACCAAGCCTAGCAAGGTGAAGGCGTTTGGGCCGGGGCTGCAGGGAGGCAGTGCGGGCTCCCCCGCC
CGCTTCACCATCGACACCAAGGGCGCCGGCACAGGTGGCCTGGGCCTGACGGTGGAGGGCCCCTGTGAGG
CGCAGCTCGAGTGCTTGGACAATGGGGATGGCACATGTTCCGTGTCCTACGTGCCCACCGAGCCCGGGGA
CTACAACATCAACATCCTCTTCGCTGACACCCACATCCCTGGCTCCCCATTCAAGGCCCACGTGGTTCCC
TGCTTTGACGCATCCAAAGTCAAGTGCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTGGGGAGGTGGGCC
AATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGCGGAGCTGACCATTGAGATCTGCTCGGAGGCGGGGCT
TCCGGCCGAGGTGTACATCCAGGACCACGGTGATGGCACGCACACCATTACCTACATTCCCCTCTGCCCC
GGGGCCTACACCGTCACCATCAAGTACGGCGGCCAGCCCGTGCCCAACTTCCCCAGCAAGCTGCAGGTGG
AACCTGCGGTGGACACTTCCGGTGTCCAGTGCTATGGGCCTGGTATTGAGGGCCAGGGTGTCTTCCGTGA
GGCCACCACTGAGTTCAGTGTGGACGCCCGGGCTCTGACACAGACCGGAGGGCCGCACGTCAAGGCCCGT
GTGGCCAACCCCTCAGGCAACCTGACGGAGACCTACGTTCAGGACCGTGGCGATGGCATGTACAAAGTGG
AGTACACGCCTTACGAGGAGGGACTGCACTCCGTGGACGTGACCTATGACGGCAGTCCCGTGCCCAGCAG
CCCCTTCCAGGTGCCCGTGACCGAGGGCTGCGACCCCTCCCGGGTGCGTGTCCACGGGCCAGGCATCCAA
AGTGGCACCACCAACAAGCCCAACAAGTTCACTGTGGAGACCAGGGGAGCTGGCACGGGCGGCCTGGGCC
TGGCTGTAGAGGGCCCCTCCGAGGCCAAGATGTCCTGCATGGATAACAAGGACGGCAGCTGCTCGGTCGA
GTACATCCCTTATGAGGCTGGCACCTACAGCCTCAACGTCACCTATGGTGGCCATCAAGTGCCAGGCAGT
CCTTTCAAGGTCCCTGTGCATGATGTGACAGATGCGTCCAAGGTCAAGTGCTCTGGGCCCGGCCTGAGCC
CAGGCATGGTTCGTGCCAACCTCCCTCAGTCCTTCCAGGTGGACACAAGCAAGGCTGGTGTGGCCCCATT
GCAGGTCAAAGTGCAAGGGCCCAAAGGCCTGGTGGAGCCAGTGGACGTGGTAGACAACGCTGATGGCACC
CAGACCGTCAATTATGTGCCCAGCCGAGAAGGGCCCTACAGCATCTCAGTACTGTATGGAGATGAAGAGG
TACCCCGGAGCCCCTTCAAGGTCAAGGTGCTGCCTACTCATGATGCCAGCAAGGTGAAGGCCAGTGGCCC
CGGGCTCAACACCACTGGCGTGCCTGCCAGCCTGCCCGTGGAGTTCACCATCGATGCAAAGGACGCCGGG
GAGGGCCTGCTGGCTGTCCAGATCACGGATCCCGAAGGCAAGCCGAAGAAGACACACATCCAAGACAACC
ATGACGGCACGTATACAGTGGCCTACGTGCCAGACGTGACAGGTCGCTACACCATCCTCATCAAGTACGG
TGGTGACGAGATCCCCTTCTCCCCGTACCGCGTGCGTGCCGTGCCCACCGGGGACGCCAGCAAGTGCACT
GTCACAGTGTCAATCGGAGGTCACGGGCTAGGTGCTGGCATCGGCCCCACCATTCAGATTGGGGAGGAGA
CGGTGATCACTGTGGACACTAAGGCGGCAGGCAAAGGCAAAGTGACGTGCACCGTGTGCACGCCTGATGG
CTCAGAGGTGGATGTGGACGTGGTGGAGAATGAGGACGGCACTTTCGACATCTTCTACACGGCCCCCCAG
CCGGGCAAATACGTCATCTGTGTGCGCTTTGGTGGCGAGCACGTGCCCAACAGCCCCTTCCAAGTGACGG
CTCTGGCTGGGGACCAGCCCTCGGTGCAGCCCCCTCTACGGTCTCAGCAGCTGGCCCCACAGTACACCTA
CGCCCAGGGCGGCCAGCAGACTTGGGCCCCGGAGAGGCCCCTGGTGGGTGTCAATGGGCTGGATGTGACC
AGCCTGAGGCCCTTTGACCTTGTCATCCCCTTCACCATCAAGAAGGGCGAGATCACAGGGGAGGTTCGGA
TGCCCTCAGGCAAGGTGGCGCAGCCCACCATCACTGACAACAAAGACGGCACCGTGACCGTGCGGTATGC
ACCCAGCGAGGCTGGCCTGCACGAGATGGACATCCGCTATGACAACATGCACATCCCAGGAAGCCCCTTG
CAGTTCTATGTGGATTACGTCAACTGTGGCCATGTCACTGCCTATGGGCCTGGCCTCACCCATGGAGTAG
TGAACAAGCCTGCCACCTTCACCGTCAACACCAAGGATGCAGGAGAGGGGGGCCTGTCTCTGGCCATTGA
GGGCCCGTCCAAAGCAGAAATCAGCTGCACTGACAACCAGGATGGGACATGCAGCGTGTCCTACCTGCCT
GTGCTGCCGGGGGACTACAGCATTCTAGTCAAGTACAATGAACAGCACGTCCCAGGCAGCCCCTTCACTG
CTCGGGTCACAGGTGACGACTCCATGCGTATGTCCCACCTAAAGGTCGGCTCTGCTGCCGACATCCCCAT
CAACATCTCAGAGACGGATCTCAGCCTGCTGACGGCCACTGTGGTCCCGCCCTCGGGCCGGGAGGAGCCC
TGTTTGCTGAAGCGGCTGCGTAATGGCCACGTGGGGATTTCATTCGTGCCCAAGGAGACGGGGGAGCACC
TGGTGCATGTGAAGAAAAATGGCCAGCACGTGGCCAGCAGCCCCATCCCGGTGGTGATCAGCCAGTCGGA
AATTGGGGATGCCAGTCGTGTTCGGGTCTCTGGTCAGGGCCTTCACGAAGGCCACACCTTTGAGCCTGCA
GAGTTTATCATTGATACCCGCGATGCAGGCTATGGTGGGCTCAGCCTGTCCATTGAGGGCCCAGCAAGG
TGGACATCAACACAGAGGACCTGGAGGACGGGACGTGCAGGGTCACCTACTGCCCCACAGAGCCAGGCAA
CTACATCATCAACATCAAGTTTGCCGACCAGCACGTGCCTGGCAGCCCCTTCTCTGTGAAGGTGACAGGC
GAGGGCCGGGTGAAAGAGAGCATCACCCGCAGGCGTCGGGCTCCTTCAGTGGCCAACGTTGGTAGTCATT
GTGACCTCAGCCTGAAAATCCCTGAAATTAGCATCCAGGATATGACAGCCCAGGTGACCAGCCCATCGGG
CAAGACCCATGAGGCCGAGATCGTGGAAGGGGAGAACCACACCTACTGCATCCGCTTTGTTCCCGCTGAG
ATGGGCACACACAGTCAGCGTCAAGTACAAGGGCCAGCACGTGCCTGGGAGCCCCTTCCAGTTCACCG
TGGGGCCCCTAGGGGAAGGGGGAGCCCACAAGGTCCGAGCTGGGGCCCTGGCCTGGAGAGAGCTGAAGC
TGGAGTGCCAGCCGAATTCAGTATCTGGACCCGGGAAGCTGGTGCTGGAGGCCTGGCCATTGCTGTCGAG
GGCCCCAGCAAGGCTGAGATCTCTTTTGAGGACCGCAAGGACGGCTCCTGTGGTGTGGCTTATGTGGTCC
AGGAGCCAGGTGACTACGAAGTCTCAGTCAAGTTCAACGAGGAACACATTCCCGACAGCCCCTTCGTGGT
GCCTGTGGCTTCTCCGTCTGGCGACGCCCGCCGCCTCACTGTTTCTAGCCTTCAGGAGTCAGGGCTAAAG
GTCAACCAGCCAGCCTCTTTTGCAGTCAGCCTGAACGGGGCCAAGGGGGCGATCGATGCCAAGGTGCACA
GCCCCTCAGGAGCCCTGGAGGAGTGCTATGTCACAGAAATTGACCAAGATAAGTATGCTGTGCGCTTCAT
CCCTCGGGAGAATGGCGTTTACCTGATTGACGTCAAGTTCAACGGTACCCACATCCCTGGAAGCCCCTTC
```

104

```
AAGATCCGAGTTGGGGAGCCTGGGCATGGAGGGGACCCAGGCTTGGTGTCTGCTTACGGAGCAGGTCTGG
AAGGCGGTGTCACAGGGAACCCAGCTGAGTTCGTCGTGAACACGAGCAATGCGGGAGCTGGTGCCCTGTC
GGTGACCATTGACGGCCCCTCCAAGGTGAAGATGGATTGCCAGGAGTGCCCTGAGGGCTACCGCGTCACC
TATACCCCCATGGCACCTGGCAGCTACCTCATCTCCATCAAGTACGGCGGCCCCTACCACATTGGGGGCA
GCCCCTTCAAGGCCAAAGTCACAGGCCCCGTCTCGTCAGCAACCACAGCCTCCACGAGACATCATCAGT
GTTTGTAGACTCTCTGACCAAGGCCACCTGTGCCCCCCAGCATGGGGCCCCGGGTCCTGGGCCTGCTGAC
GCCAGCAAGGTGGTGGCCAAGGGCCTGGGGCTGAGCAAGGCCTACGTAGGCCAGAAGAGCAGCTTCACAG
TAGACTGCAGCAAAGCAGGCAACAACATGCTGCTGGTGGGGGTTCATGGCCCAAGGACCCCCTGCGAGGA
GATCCTGGTGAAGCACGTGGGCAGCCGGCTCTACAGCGTGTCCTACCTGCTCAAGGACAAGGGGGAGTAC
ACACTGGTGGTCAAATGGGGGCACGAGCACATCCCAGGCAGCCCCTACCGCGTTGTGGTGCCCTGAGTCT
GGGGCCCGTGCCAGCCGGCAGCCCCCAAGCCTGCCCCGCTACCCAAGCAGCCCCGCCCTCTTCCCCTCAA
CCCCGGCCCAGGCCGCCCTGGCCGCCCGCCTGTCACTGCAGCTGCCCCCTGCCCTGTGCCGTGCTGCGCTC
ACCTGCCTCCCCAGCCAGCCGCTGACCTCTCGGCTTTCACTTGGGCAGAGGGAGCCATTTGGTGGCGCTG
CTTGTCTTCTTTGGTTCTGGGAGGGGTGAGGGATGGGG
```

>gi|4503745|ref|NP_001447.1| filamin 1 (actin-binding protein-280); filamin
A, alpha (actin-binding protein-280); filamin 1; actin-binding protein-280
[Homo sapiens] (SEQ ID NO:84)

```
MSSSHSRAGQSAAGAAPGGGVDTRDAEMPATEKDLAEDAPWKKIQQNTFTRWCNEHLKCVSKRIANLQTD
LSDGLRLIALLEVLSQKKMHRKHNQRPTFRQMQLENVSVALEFLDRESIKLVSIDSKAIVDGNLKLILGL
IWTLILHYSISMPMWDEEEDEEAKKQTPKQRLLGWIQNKLPQLPITNFSRDWQSGRALGALVDSCAPGLC
PDWDSWDASKPVTNAREAMQQADDWLGIPQVITPEEIVDPNVDEHSVMTYLSQFPKAKLKPGAPLRPKLN
PKKARAYGPGIEPTGNMVKKRAEFTVETRSAGQGEVLVYVEDPAGHQEEAKVTANNDKNRTFSVWYVPEV
TGTHKVTVLFAGQHIAKSPFEVYVDKSQGDASKVTAQGPGLEPSGNIANKTTYFEIFTAGAGTGEVEVVI
QDPMGQKGTVEPQLEARGDSTYRCSYQPTMEGVHTVHVTFAGVPIPRSPYTVTVGQACNPSACRAVGRGL
QPKGVRVKETADFKVYTKGAGSELKVTVKGPKGEERVKQKDLGDGVYGFEYYPMVPGTYIVTITWGGQN
IGRSPFEVKVGTECGNQKVRAWGPGLEGGVVGKSADFVVEAIGDDVGTLGFSVEGPSQAKIECDDKGDGS
CDVRYWPQEAGEYAVHVLCNSEDIRLSPFMADIRDAPQDFHPDRVKARGPGLEKTGVAVNKPAEFTVDAK
HGGKAPLRVQVQDNEGCPVEALVKDNGNGTYSCSYVPRKPVKHTAMVSWGGVSIPNSPFRVNVGAGSHPN
KVKVYGPGVAKTGLKAHEPTYFTVDCAEAGQGDVSIGIKCAPGVVGPAEADIDFDIIRNDNDTFTVKYTP
RGAGSYTIMVLFADQATPTSPIRVKVEPSHDASKVKAEGPGLSRTGVELGKPTHFTVNAKAAGKGKLDVQ
FSGLTKGDAVRDVDIIDHHDNTYTVKYTPVQQGPVGVNVTYGGDPIPKSPFSVAVSPSLDLSKIKVSGLG
EKVDVGKDQEFTVKSKGAGGQGKVASKIVGPSGAAVPCKVEPGLGADNSVVRFLPREEGPYEVEVTYDGV
PVPGSPFPLEAVAPTKPSKVKAFGPGLQGGSAGSPARFTIDTKGAGTGGLGLTVEGPCEAQLECLDNGDG
TCSVSYVPTEPGDYNINILFADTHIPGSPFKAHVVPCFDASKVKCSGPGLERATAGEVGQFQVDCSSAGS
AELTIEICSEAGLPAEVYIQDHGDGTHTITYIPLCPGAYTVTIKYGGQPVPNFPSKLQVEPAVDTSGVQC
YGPGIEGQGVFREATTEFSVDARALTQTGGPHVKARVANPSGNLTETYVQDRGDGMYKVEYTPYEEGLHS
VDVTYDGSPVPSSPFQVPVTEGCDPSRVRVHGPGIQSGTTNKPNKFTVETRGAGTGGLGLAVEGPSEAKM
SCMDNKDGSCSVEYIPYEAGTYSLNVTYGGHQVPGSPFKVPVHDVTDASKVKCSGPGLSPGMVRANLPQS
FQVDTSKAGVAPLQVKVQGPKGLVEPVDVVDNADGTQTVNYVPSREGPYSISVLYGDEEVPRSPFKVKVL
PTHDASKVKASGPGLNTTGVPASLPVEFTIDAKDAGEGLLAVQITDPEGKPKKTHIQDNHDGTYTVAYVP
DVTGRYTILIKYGGDEIPFSPYRVRAVPTGDASKCTVTVSIGGHGLGAGIGPTIQIGEETVITVDTKAAG
KGKVTCTVCTPDGSEVDVDVVENEDGTFDIFYTAPQPGKYVICVRFGGEHVPNSPFQVTALAGDQPSVQP
PLRSQQLAPQYTYAQGGQQTWAPERPLVGVNGLDVTSLRPFDLVIPFTIKKGEITGEVRMPSGKVAQPTI
TDNKDGTVTVRYAPSEAGLHEMDIRYDNMHIPGSPLQFYVDYVNCGHVTAYGPGLTHGVVNKPATFTVNT
KDAGEGGLSLAIEGPSKAEISCTDNQDGTCSVSYLPVLPGDYSILVKYNEQHVPGSPFTARVTGDDSMRM
SHLKVGSAADIPINISETDLSLLTATVVPPSGREEPCLLKRLRNGHVGISFVPKETGEHLVHVKKNGQHV
ASSPIPVVISQSEIGDASRVRVSGQGLHEGHTFEPAEFIIDTRDAGYGGLSLSIEGPSKVDINTEDLEDG
TCRVTYCPTEPGNYIINIKFADQHVPGSPFSVKVTGEGRVKESITRRRRAPSVANVGSHCDLSLKIPEIS
IQDMTAQVTSPSGKTHEAEIVEGENHTYCIRFVPAEMGTHTVSVKYKGQHVPGSPFQFTVGPLGEGGAHK
VRAGGPGLERAEAGVPAEFSIWTREAGAGGLAIAVEGPSKAEISFEDRKDGSCGVAYVVQEPGDYEVSVK
FNEEHIPDSPFVVPVASPSGDARRLTVSSLQESGLKVNQPASFAVSLNGAKGAIDAKVHSPSGALEECYV
TEIDQDKYAVRFIPRENGVYLIDVKFNGTHIPGSPFKIRVGEPGHGGDPGLVSAYGAGLEGGVTGNPAEF
VVNTSNAGAGALSVTIDGPSKVKMDCQECPEGYRVTYTPMAPGSYLISIKYGGPYHIGGSPFKAKVTGPR
LVSNHSLHETSSVFVDSLTKATCAPQHGAPGPGPADASKVVAKGLGLSKAYVGQKSSFTVDCSKAGNNML
LVGVHGPRTPCEEILVKHVGSRLYSVSYLLKDKGEYTLVVKWGHEHIPGSPYRVVVP
```

centractin alpha

>gi|5031568|ref|NM_005736.1| Homo sapiens ARP1 (actin-related protein 1,
yeast) homolog A (centractin alpha) (ACTR1A), mRNA (SEQ ID NO:85)
GAATTCGGGGCGCTACGGCGGACCCGGCTGGGCAGTTCCTTCCCCAGAAGGAGAGATTCCTCTGCCATGG
AGTCCTACGATGTGATCGCCAACCAGCCTGTCGTGATCGACAACGGATCCGGTGTGATTAAAGCTGGTTT
TGCTGGTGATCAGATCCCCAAATACTGCTTTCCAAACTATGTGGGCCGACCCAAGCACGTTCGTGTCATG
GCAGGAGCCCTTGAAGGCGACATCTTCATTGGCCCCAAAGCTGAGGAGCACCGAGGGCTGCTTTCAATCC
GCTATCCCATGGAGCATGGCATCGTCAAGGATTGGAACGACATGGAACGCATTTGGCAATATGTCTATTC
TAAGGACCAGCTGCAGACTTTCTCAGAGGAGCATCCTGTGCTCCTGACTGAGGCGCCTTTAAACCCACGA
AAAAACCGGGAACGAGCTGCCGAAGTTTTCTTCGAGACCTTCAATGTGCCCGCTCTTTTCATCTCCATGC
AAGCTGTACTCAGCCTTTACGCTACAGGCAGGACCACAGGGGTGGTGCTGGATTCTGGGGATGGAGTCAC
CCATGCTGTGCCCATCTATGAGGGCTTTGCCATGCCCCACTCCATCATGCGCATCGACATCGCGGGCCGG
GACGTCTCTCGCTTCCTGCGCCTCTACCTGCGTAAGGAGGGCTACGACTTCCACTCATCCTCTGAGTTTG
AGATTGTCAAGGCCATAAAAGAAAGAGCCTGTTACCTATCCATAAACCCCCAAAAGGATGAGACGCTAGA
GACAGAGAAAGCTCAGTACTACCTGCCTGATGGCAGCACCATTGAGATTGGTCCTTCCCGATTCCGGGCC
CCTGAGTTGCTCTTCAGGCCAGATTTGATTGGAGAGGAGAGTGAAGGCATCCACGAGGTCCTGGTGTTCG
CCATTCAGAAGTCAGACATGGACCTGCGGCGCACGCTTTTCTCTAACATTGTCCTCTCAGGAGGCTCTAC
CCTGTTCAAAGGTTTTGGTGACAGGCTCCTGAGTGAAGTGAAGAAACTAGCTCCAAAAGATGTGAAGATC
AGGATATCTGCACCTCAGGAGAGACTGTATTCCACGTGGATTGGGGGCTCCATCCTTGCCTCCCTGGACA
CCTTTAAGAAGATGTGGGTCTCCAAAAAGGAATATGAGGAAGACGGTGCCCGATCCATCCACAGAAAAAC
CTTCTAATGTCGGGACATCATCTTCACCTCTCTCTGAAGTTAACTCCACTTTAAAACTCGCTTTCTTGAG
TCGGAGTGTTTGCGAGGAACTGCCTGTGTGTGAGTGCGTGTGTGGATATGAGTGTGTGCGCACATGCGAG
TGCCGTGTGGCCCTGGGACCCTGGGCCCAGAAAGGACGATGAACTACCCGCAGTGGTGATGCCTGAGGCC
TGGGGTTGACCACTAACTGGCTCCTGACAGGGAAGAGCGCTGGCAGAGGCTGTGCTCCCTCCTCAGGTGG
CCTCTGGCTGGCTGTGGGGGACTCCGTTTACTACCACAGGGAGACAGAGGGAGGTAAGCCATCCCCCGGG
AGACCTTGCTGCTGACCATCCTAGGCTGGGCTGGCCCACCCTCACCCCCACCCCCAGGGTGCCCTGAGGC
CCCAGGCAGCTGCTGCCTCCACTATCGATGCCTCCTGACTGCACACTGAGGACTGGGACTGGGGTTGAGT
TCTGTCTGGTTTTGTTGCCATTTTGGTTTGGGAGGCTGGAAAAGCACCCCAAGAAGCTATTACAGAGACT
GGAGTCAGGAGAGAGCAGGAGGCCCTCATGTTCACCAGGGAACAGGACCACACCGGCCACTGAAGGAGGG
CAGGAGCAGTCCTCCCTCTGAATGGCTGCAGAGTTAATGTTCCCAGCCCAGTCCCCTTTCGGGGGCCTTG
GGAGAGTTTAAGGCACCTGCTGGTTCCAGGACCTCGCTTTCCATCTGTTCTTGTTGCAATGCCATCTTCA
AACCGTTTTATTTATTGAAGTGTTTGTTCAGTTAGGGGCTGGAGAGAGGGAGCTTGCTGCCTCCTGCCTT
GCTACACTAATGTTTACAGCACCTAAGCTTAGCCTCCAGGGCCCCACCTCTCCCAGCTGATGGTGAGCTG
ACAGTGTCCACAGGTTCCAGGACCATTTGAGATTGGAAGCTACACTCAAAGACACTCCACCAGGCTCTT
TCTCCCTTTTCCTCTTCTCACTGCCCTGGAATCAACAGGCTGGTTGCTGGTTAGATTTTCTGAAACAGGA
GGTAAAATTTTTCTTTGGCAGAGGCCCCTAAGCAAGGGAGGGGTGTTGGAGAGCCAGTGCCCTTAAGACT
GGAGAAAGCTGCAATTTACCAAGTTGCCTTTTGCCACTGTAGCTGACCAGGGGACTAGGTTGTAGAGGTG
GGAAGGCCCCTCTGGGCTGATCTTGTGCCATTCTTGACCTTGGACCTGCTTGGTTAAGGAGGGAGTGGGC
CAGACCAGAGTGCCAGGAGCTAATGGAGCCAGGCCTGACACCTAGGAGTGGTCCAAAGCCTTCAGCCTAG
ATGGTGCAAAGCTGGGGCCAGCCTGTCTTCACCGGCACCCTCACCTGTGACACCAAGACCCACCCCAATC
CAGACTTCACACAGTATTCTCCCCCACGCCGTCTATGACCAAAGGCCCCTGCCAGGTGTGGGTCCACAGC
AGCAGGTATGTGTGAAAGCAACGTAGCGCCCCGCGGACTGCAGTGCGCTTAACCAACTCACCTCCCTTCT
CTTAGCCCAAGCCTGTCCCTCGCACAGCCTCGCACAAACCACATTGCCTGGTGGGGCCCAGTGTACTGAA
ATAAAGTCGTTCCGATAGACACGTCAAAAAAAAAAAACCCGAATTC

>gi|5031569|ref|NP_005727.1| ARP1 actin-related protein 1 homolog A,
centractin alpha; ARP1 (actin-related protein 1, yeast) homolog A
(centractin alpha); centractin alpha; actin-RPV; centrosome-associated
actin homolog; ARP1, yeast homolog A [Homo sapiens] (SEQ ID NO:86)
MESYDVIANQPVVIDNGSGVIKAGFAGDQIPKYCFPNYVGRPKHVRVMAGALEGDIFIGPKAEEHRGLLS
IRYPMEHGIVKDWNDMERIWQYVYSKDQLQTFSEEHPVLLTEAPLNPRKNRERAAEVFFETFNVPALFIS
MQAVLSLYATGRTTGVVLDSGDGVTHAVPIYEGFAMPHSIMRIDIAGRDVSRFLRLYLRKEGYDFHSSSE
FEIVKAIKERACYLSINPQKDETLETEKAQYYLPDGSTIEIGPSRFRAPELLFRPDLIGEESEGIHEVLV
FAIQKSDMDLRRTLFSNIVLSGGSTLFKGFGDRLLSEVKKLAPKDVKIRISAPQERLYSTWIGGSILASL
DTFKKMWVSKKEYEEDGARSIHRKTF

moesin

>gi|4505256|ref|NM_002444.1| Homo sapiens moesin (MSN), mRNA (SEQ ID NO:87)
GGCACGAGGCCAGCCGAATCCAAGCCGTGTGTACTGCGTGCTCAGCACTGCCCGACAGTCCTAGCTAAAC
TTCGCCAACTCCGCTGCCTTTGCCGCCACCATGCCCAAAACGATCAGTGTGCGTGTGACCACCATGGATG

```
CAGAGCTGGAGTTTGCCATCCAGCCCAACACCACCGGGAAGCAGCTATTTGACCAGGTGGTGAAAACTAT
TGGCTTGAGGGAAGTTTGGTTCTTTGGTCTGCAGTACCAGGACACTAAAGGTTTCTCCACCTGGCTGAAA
CTCAATAAGAAGGTGACTGCCCAGGATGTGCGGAAGGAAAGCCCCCTGCTCTTTAAGTTCCGTGCCAAGT
TCTACCCTGAGGATGTGTCCGAGGAATTGATTCAGGACATCACTCAGCGCCTGTTCTTTCTGCAAGTGAA
```
5
```
AGAGGGCATTCTCAATGATGATATTTACTGCCCGCCTGAGACCGCTGTGCTGCTGGCCTCGTATGCTGTC
CAGTCTAAGTATGGCGACTTCAATAAGGAAGTGCATAAGTCTGGCTACCTGGCCGGAGACAAGTTGCTCC
CGCAGAGAGTCCTGGAACAGCACAAACTCAACAAGGACCAGTGGGAGGAGCGGATCCAGGTGTGGCATGA
GGAACACCGTGGCATGCTCAGGGAGGATGCTGTCCTGGAATATCTGAAGATTGCTCAAGATCTGGAGATG
TATGGTGTGAACTACTTCAGCATCAAGAACAAGAAAGGCTCAGAGCTGTGGCTGGGGGTGGATGCCCTGG
```
10
```
GTCTCAACATCTATGAGCAGAATGACAGACTAACTCCCAAGATAGGCTTCCCCTGGAGTGAAATCAGGAA
CATCTCTTTCAATGATAAGAAATTTGTCATCAAGCCCATTGACAAAAAGCCCCGGACTTCGTCTTCTAT
GCTCCCCGGCTGCGGATTAACAAGCGGATCTTGGCCTTGTGCATGGGGAACCATGAACTATACATGCGCC
GTCGCAAGCCTGATACCATTGAGGTGCAGCAGATGAAGGCACAGGCCCGGGAGGAGAAGCACCAGAAGCA
GATGGAGCGTGCTATGCTGGAAAATGAGAAGAAGAAGCGTGAAATGGCAGAGAAGGAGAAAGAGAAGATT
```
15
```
GAACGGGAGAAGGAGGAGCTGATGGAGAGGCTGAAGCAGATCGAGGAACAGACTAAGAAGGCTCAGCAAG
AACTGGAAGAACAGACCCGTAGGGCTCTGGAACTTGAGCAGGAACGGAAGCGTGCCCAGAGCGAGGCTGA
AAAGCTGGCCAAGGAGCGTCAAGAAGCTGAAGAGGCCAAGGAGGCCTTGCTGCAGGCCTCCCGGGACCAG
AAAAAGACTCAGGAACAGCTGGCCTTGGAAATGGCAGAGCTGACAGCTCGAATCTCCCAGCTGGAGATGG
CCCGACAGAAGAAGGAGAGTGAGGCTGTGGAGTGGCAGCAGAAGGCCCAGATGGTACAGGAAGACTTGGA
```
20
```
GAAGACCCGTGCTGAGCTGAAGACTGCCATGAGTACACCTCATGTGGCAGAGCCTGCTGAGAATGAGCAG
GATGAGCAGGATGAGAATGGGGCAGAGGCTAGTGCTGACCTACGGGCTGATGCTATGGCCAAGGACCGCA
GTGAGGAGGAACGTACCACTGAGGCAGAGAAGAATGAGCGTGTGCAGAAGCACCTGAAGGCCCTCACTTC
GGAGCTGGCCAATGCCAGAGATGAGTCCAAGAAGACTGCCAATGACATGATCCATGCTGAGAACATGCGA
CTGGGCCGAGACAAATACAAGACCCTGCGCCAGATCCGGCAGGGCAACACCAAGCAGCGCATTGACGAAT
```
25
```
TTGAGTCTATGTAATGGGCACCCAGCCTCTAGGGACCCCTCCTCCCTTTTTCCTTGTCCCACACTCCTA
CACCTAACTCACCTAACTCATACTGTGCTGGAGCCACTAACTAGAGCAGCCCTGGAGTCATGCCAAGCAT
TTAATGTAGCCATGGGACCAAACCTAGCCCCTTAGCCCCCACCCACTTCCCTGGGCAAATGAATGGCTCA
CTATGGTGCCAATGGAACCTCCTTTCTCTTCTCTGTTCCATTGAATCTGTATGGCTAGAATATCCTACTT
CTCCAGCCTAGAGGTACTTTCCACTTGATTTTGCAAATGCCCTTACACTTACTGTTGTCCTATGGGAGTC
```
30
```
AAGTGTGGAGTAGGTTGGAAGCTAGCTCCCCTCCTCTCCCCTCCACTGTCTTCTTCAGGTCCTGAGATTA
CACGGTGGAGTGTATGCGGTCTAGGAATGAGACAGGACCTAGATATCTTCTCCAGGGATGTCAACTGACC
TAAAATTTGCCCTCCCATCCCGTTTAGAGTTATTTAGGCTTTGTAACGATTGGGGGAATAAAAAGATGTT
CAGTCATTTTTGTTTCTACCTCCCAGATCGGATCTGTTGCAAACTCAGCCTCAATAAGCCTTGTCGTTGA
CTTTAGGGACTCAATTTCTCCCCAGGGTGGATGGGGGAAATGGTGCCTTCAAGACCTTCACCAAACATAC
```
35
```
TAGAAGGGCATTGGCCATTCTATTGTGGCAAGGCTGAGTAGAAGATCCTACCCCAATTCCTTGTAGGAGT
ATAGGCCGGTCTAAAGTGAGCTCTATGGGCAGATCTACCCCTTACTTATTATTCCAGATCTGCAGTCACT
TCGTGGGATCTGCCCCTCCCTGCTTCAATACCCAAATCCTCTCCAGCTATAACAGTAGGGATGAGTACCC
AAAAGCTCAGCCAGCCCCATCAGGACTCTTGTGAAAGAGAGGATATGTTCACACCTAGCGTCAGTATTT
TCCCTGCTAGGGGTTTTAGGTCTCTTCCCCTCTCAGAGCTACTTGGGCCATAGCTCCTGCTCCACAGCCA
```
40
```
TCCCAGCCTTGGCATCTAGAGCTTGATGCCAGTAGGCTCAACTAGGGAGTGAGTGCAAAAAGCTGAGTAT
GGTGAGAGAAGCCTGTGCCCTGATCCAAGTTTACTCAACCCTCTCAGGTGACCAAAATCCCCTTCTCATC
ACTCCCCTCAAAGAGGTGACTGGGCCCTGCCTCTGTTTGACAAACCTCTAACCCAGGTCTTGACACCAGC
TGTTCTGTCCCTTGGAGCTGTAAACCAGAGAGCTGCTGGGGGATTCTGGCCTAGTCCCTTCCACACCCCC
ACCCCTTGCTCTCAACCCAGGAGCATCCACCTCCTTCTCTGTCTCATGTGTGCTCTTCTTCTTTCTACAG
```
45
```
TATTATGTACTCTACTGATATCTAAATATTGATTTCTGCCTTCCTTGCTAATGCACCATTAGAAGATATT
AGTCTTGGGGCAGGATGATTTTGGCCTCATTACTTTACCACCCCACACCTGGAAAGCATATACTATATT
ACAAATGACATTTTGCCAAAATTATTAATATAAGAAGCTTTCAGTATTAGTGATGTCATCTGTCACTAT
AGGTCATACAATCCATTCTTAAAGTACTTGTTATTTGTTTTTATTATTACTGTTTGTCTTCTCCCCAGGG
TTCAGTCCCTCAAGGGGCCATCCTGTCCCACCATGCAGTGCCCCCTAGCTTAGAGCCTCCCTCAATTCCC
```
50
```
CCTGGCCACCACCCCCACTCTGTGCCTGACCTTGAGGAGTCTTGTGTGCATTGCTGTGAATTAGCTCAC
TTGGTGATATGTCCTATATTGGCTAAATTGAAACCTGGAATTGTGGGGCAATCTATTAATAGCTGCCTTA
AAGTCAGTAACTTACCCTTAGGGAGGCTGGGGGAAAAGGTTAGATTTTGTATTCAGGGGTTTTTTGTGTA
CTTTTTGGGTTTTTAAAAAATTGTTTTTGGAGGGGTTTATGCTCAATCCATGTTCTATTTCAGTGCCAAT
AAAATTTAGGTGACTTCAAAAAAAAAAAA
```
55
```
>gi|4505257|ref|NP_002435.1| moesin [Homo sapiens] (SEQ ID NO:88)
MPKTISVRVTTMDAELEFAIQPNTTGKQLFDQVVKTIGLREVWFFGLQYQDTKGFSTWLKLNKKVTAQDV
RKESPLLFKFRAKFYPEDVSEELIQDITQRLFFLQVKEGILNDDIYCPPETAVLLASYAVQSKYGDFNKE
VHKSGYLAGDKLLPQRVLEQHKLNKDQWEERIQVWHEEHRGMLREDAVLEYLKIAQDLEMYGVNYFSIKN
```
60
```
KKGSELWLGVDALGLNIYEQNDRLTPKIGFPWSEIRNISFNDKKFVIKPIDKKAPDFVFYAPRLRINKRI
LALCMGNHELYMRRRKPDTIEVQQMKAQAREEKHQKQMERAMLENEKKKREMAEKEKEKIEREKEELMER
```

LKQIEEQTKKAQQELEEQTRRALELEQERKRAQSEAEKLAKERQEAEEAKEALLQASRDQKKTQEQLALE
MAELTARISQLEMARQKKESEAVEWQQKAQMVQEDLEKTRAELKTAMSTPHVAEPAENEQDEQDENGAEA
SADLRADAMAKDRSEEERTTEAEKNERVQKHLKALTSELANARDESKKTANDMIHAENMRLGRDKYKTLR
QIRQGNTKQRIDEFESM

5

TIMP3

>gi|9257248|ref|NM_000362.2| Homo sapiens tissue inhibitor of
10 metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory) (TIMP3),
mRNA (SEQ ID NO:89)
TCTGTCGACTTGCCCCAGAGCTGATCCTTGTCTTTGTCCACTTCTCAGCGAGGATGGCACTTCAGGGAGC
CCTTCCCTTACTATCGCAGAGAGAGCAGGCCCTCCCCAGTCATGTCCAACCCAGAACTCTGTTTTGTTTT
CTTCATAGCCCTAGCATCACAGAAAATCACCCTGTGCATTCATGGATGTCCACGGGGGCAAGGGCTTTGT
15 GTTGCTTAACCCAGCATCCTGAACCGTGTTTGTTGAATGAATACAGAACCCCGTTTGCTCTGGGAGAGCA
CAGAAAACAGTCTTCTATCATATATCATAGCCAGCTCGAAACAGCAGATGGCTTCCATATCCAGAGAGCA
AGAACCAGAGAGAGAGAGAAAGAGAGAGAGTTTGGGTCTTTCTCCTCTNTNCCTGCTCTCTCCAGAGAAA
CTGGAGGGGTAGCAGTTAGCATTCCCCGCTGGTTCCACCAAGCACAGTCAAGGTCTCTAGGACATGGCC
ACCCCTCACCTGTGGAAGCGGTCCTGCTGGGTGGGTGGGTGTTAGTTGGTTCTGGTTTGGGTCAGAGAC
20 ACCCAGTGGCCCAGGTGGGCGTGGGGCCAGGGCGCAGACGAGAAGGGGCACGAGGGCTCCGCTCCGAGGA
CCCAGCGGCAAGCACCGGTCCCGGGNGGGCCCCAGCCCACCCACTCGCGTGCCCACGGCGGCATTATTCC
CTATAAGGATCTGAACGATCCGGGGGCGGCCCCGCCCCGTTACCCCTTGCCCCCGGCCCCGCCCCCTTTT
TGGAGGGCCGATGAGGTAATGCGGCTCTGCCATTGGTCTGAGGGGGCGGGCCCCAACAGCCCGAGGCGGG
GTCCCCGGGGGCCCAGCGCTATATCACTCGGCCGCCCAGGCAGACGGCGCAGAGCGGGCAGCAGGCAGGC
25 GGCGGGCGCTCAGACGGCTTCTCCTCCTCCTCTTGCTCCTCCAGCTCCTGCTCCTTCGCCGGGAGGCCGC
CCGCCGAGTCCTGGCCAGCGCCGAGGCAGCCTGCCTGCGCCCCATCCCGTCCCGCCGGGCACTCGGAGGG
CAGCGCGCCGGAGGCCAAGGTTGCCCCGCACGGCCCGGCGGGCGAGCGAGCTCGGGCTGCAGCAGCCCCG
CCGGCGCGCACGGCAACTTTGGAGAGGCGAGCAGCAGCCCCGGCAGCGGCGGCAGCAGCGGCA<u>ATG</u>ACCC
CTTGGCTCGGGCTCATCGTGCTCCTGGGCAGCTGGAGCCTGGGGGACTGGGGCGCCGAGGCGTGCACATG
30 CTCGCCCAGCCACCCCCAGGACGCCTTCTGCAACTCCGACATCGTGATCCGGGCCAAGGTGGTGGGGAAG
AAGCTGGTAAAGGAGGGGCCCTTCGGCACGCTGGTCTACACCATCAAGCAGATGAAGATGTACCGAGGCT
TCACCAAGATGCCCCATGTGCAGTACATCCACACGGAAGCTTCCGAGAGTCTCTGTGGCCTTAAGCTGGA
GGTCAACAAGTACCAGTACCTGCTGACAGGTCGCGTCTATGATGGCAAGATGTACACGGGCTGTGCAAC
TTCGTGGAGAGGTGGGACCAGCTCACCCTCTCCCAGCGCAAGGGGCTGAACTATCGGTATCACCTGGGTT
35 GTAACTGCAAGATCAAGTCCTGCTACTACCTGCCTTGCTTTGTGACTTCCAAGAACGAGTGTCTCTGGAC
CGACATGCTCTCCAATTTCGGTTACCCTGGCTACCAGTCCAAACACTACGCCTGCATCCGGCAGAAGGGC
GGCTACTGCAGCTGGTACCGAGGATGGGCCCCCCCGGATAAAAGCATCATCAATGCCACAGACCCC<u>TGAG</u>
CGCCAGACCCTGCCCCACCTCACTTCCCTCCCTTCCCGCTGAGCTTCCCTTGGACACTAACTCTTCCCAG
ATGATGACAATGAAATTAGTGCCTGTTTTCTTGCAAATTTAGCACTTGGAACATTTAAAGAAAGGTCTAT
40 GCTGTCATATGGGGTTTATTGGGAACTATCCTCCTGGCCCCACCCTGCCCCTTCTTTTTGGTTTTGACAT
CATTCATTTCCACCTGGGAATTTCTGGTGCCATGCCAGAAAGAATGAGGAACCTGTATTCCTCTTCTTCG
TGATAATATAATCTCTATTTTTTTAGGAAAACAAAAATGAAAAACTACTCCATTTGAGGATTGTAATTCC
CACCCCTCTTGCTTCTTCCCCACCTCACCATCTCCCAGACCCTCTTCCCTTTGCCCTTCTCCTCCAATAC
ATAAAGGACACAGACAAGGAACTTGCTGAAAGGCCAACCATTTCAGGATCAGTCAAAGGCAGCAAGCAGA
45 TAGACTCAAGGTGTGTGAAAGATGTTATACACCAGGAGCTGCCACTGCATGTCCCAACCAGACTGTGTCT
GTCTGTGTCTGCATGTAAGAGTGAGGGAGGGAAGGAAGGAACTACAAGAGAGTCGGAGATGATGCAGCAC
ACACACAATTCCCCAGCCCAGTGATGCTTGTGTTGACCAGATGTTCCTGAGTCTGGAGCAAGCACCCAGG
CCAGAATAACAGAGCTTTCTTAGTTGGTGAAGACTTAAACATCTGCCTGAGGTCAGGAGGCAATTTGCCT
GCCTTGTACAAAAGCTCAGGTGAAAGACTGAGATGAATGTCTTTCCTCTCCCTGCCTCCCACCAGACTTC
50 CTCCTGGAAAACGCTTTGGTAGATTTGGCCAGGAGCTTTCTTTTATGTAAATTGGATAAATACACACACC
ATACACTATCCACAGATATAGCCAAGTAGATTTGGGTAGAGGATACTATTTCCAGAATAGTGTTTAGCTC
ACCTAGGGGGATATGTTTGTATACACATTTGCATATACCCACATGGGGACATAAGCTAATTTTTTTTACAG
GACACAGAATTCTGTTCAATGCTGTTAAATATGCCAATAGTTTAATCTCTTCTATTTTGTTGTCGTTGCT
TGTTTGAAGAAAATCATGACATTCCAAGTTGACATTTTTTTTTTCATTTTAATTAAAAATTTGAAATTCTG
55 AACACCGTCAGCACCCTCTCTTCCCTATCATGGGTCATCTGACCCCTGTCCGTCTCCTTGTCCCTGCTTC
ATGTTTGGGGGCCTTTCTTTAACTGCCTTCCTGGCTTAGCTCAGATGGCAGATGAGAGTGTAGTCAAGGG
CCTGGGCACAGGAGGGAGAGCTGCAGAGTGTCCTGCCTGCCTTGGCTGGAGGGACACCTCTCCTGGGTGT
GGAGACAGCTTGGTTCCCTTTCCCTAGCTCCCTGGTGGGTGAATGCCACCTCCTGAGATCCTCACCTCTT
GGAATTAAAATTGTTGGTCACTGGGGAAAGCCTGAGTTTGCAACCAGTTGTAGGGTTTCTGTTGTGTTTT
60 TTTTTTTTTTGAAATAAAACTATAATATAAATTCTCCTATTAAATAAAATTATTTTAAGTTTTAGTGTCA
AAAGTGAGATGCTGAGAGTAGGTGATAATGTATATTTTACAGAGTGGGGGTTGGCAGGATGGTGACATTG

```
AACATGATTGCTCTCTGTCTCTTTTTTCAGCTTATGGGTATTTATCTTCTATTAGTATTTGTATCTTCAG
TTCATTCCACTTTAGGAAACAGAGCTGCCAATTGAAACAGAAGAAGAAAAAAAAAAAGCAGCAGACAAC
ACACTGTAGAGTCTTGCACACACACAAGTGCCCAGGCAAGGTGCTTGGCAGAACCGCAGAGTGGGAAGAG
AGTACCGGCATCGGGTTTCCTTGGGATCAATTTCATTACCGTGTACCTTTCCCATTGTGGTCACGCCATT
TGGCAGGGGGAGAATGGGAGGCTTGGCCTTCTTTGTGAGGCAGTGTGAGCAGAAGAAGCTGATGCCAGCA
TGTCACTGGTTTTGAAGGGATGAGCCCAGACTTGATGTTTTGGGATTGTCCTTATTTTAACCTCAAGGTC
TCGCATGGTGGGGCCCCTGACCAACCTACACAAGTTCCCTCCCACAAGTGGACATCAGTGTCTTCTCTGT
GAGGCATCTGGCCATTCGCACTCCCTGGTGTGGTCAGCCTCTCTCACACAAGGAGGAACTTGGGTGAAGG
CTGAGTGTGAGGCACCTGAAGTTTCCCTGCGGAGTCGATAAATTAGCAGAACCACATCCCATCTGTTAG
GCCTTGGTGAGGAGGCCCTGGGCAAAGAAGGGTCTTTCGCAAAGCGATGTCAGAGGGCGGTTTTGAGCTT
TCTATAAGCTATAGCTTTGTTTATTTCACCCGTTCACTTACTGTATAATTTAAAATCATTTATGTAGCTG
AGACACTTCTGTATTTCAATCATATCATGAACATTTTATTTTGCTAAATCTTGTGTCATGTGTAGGCTGT
AATATGTGTACATTGTGTTTAAGAGAAAAATGAAACCCACATGCCGCCATTTTCCTGAATCAAATTCTGC
AGTGGAATGGAGAGGAAAATACTTCTAGGCAAGCAGCTAGACTGGTGAATTGGGGGAAATAGAAGGAACT
AGTAACTGAGACTCCTCCAGCCTCTTCCCTATTGGAATCCCAATGGCTCCTGGAGTAGGAAAAAAGTTTA
AACTACATTCATGTTCTTGTTCTGTGTCACTCGGCCCTGGGTAGTCTACCATTTACTTCACCCCAAGTCC
TGCTGCCCATCCAGTTGGGAAGCCCATGATTTTCCTAAGAATCCAGGGCCATAGGAGATACAATTCCAAG
TTCTCGCTTCCTCCTTTGGGCATCTCTTCTGCCTCCCAATCAAGGAAGCTCCACGCTCAGGCTCTCAGCT
CTCGGGCCAGTGCTCTGCTCTGTCCAGGGTAGGTAATACTGGGAGACTCCTGTCTTTTACCCTCCCCTCG
TTCCAGACCTGCCTCATGGTGGCAACATGGTTCTTGAACAATTAAAGAAACAAATGACTTTTTGGAATAG
CCCTGTCTAGGGCAAACTGTGGCCCCCAGGAGACACTACCCTTCCATGCCCCAGACCTCTGTCTTGCATG
TGACAATTGACAATCTGGACTACCCCAAGATGGCACCCAAGTGTTTGGCTTCTGGCTACCTAAGGTTAAC
ATGTCACTAGAGTATTTTTATGAGAGACAAACATTATAAAAATCTGATGGCAAAAGCAAAACAAAATGGA
AAGTAGGGGAGGTGGATGTGACAACAACTTCCAAATTGGCTCTTTGGAGGCGAGAGGAAGGGGAGAACTT
GGAGAATAGTTTTTGCTTTGGGGGTAGAGGCTTCTTAGATTCTCCCAGCATCCGCCTTTCCCTTTAGCCA
GTCTGCTGTCCTGAAACCCAGAAGTGATGGAGAGAAACCAACAAGAGATCTCGAACCCTGTCTAGAAGGA
ATGTATTTGTTGCTAAATTTCGTAGCACTGTTTACAGTTTTCCTCCATGTTATTTATGAATTTTATATTC
CGTGAATGTATATTGTCTTGTAATGTTGCATAATGTTCACTTTTTATAGTGTGTCCTTTATTCTAAACAG
TAAAGTGGTTTTATTTCTATCAC
```

>gi|4507513|ref|NP_000353.1| tissue inhibitor of metalloproteinase 3;
Tissue inhibitor of metalloproteinase-3; K222 expressed in degenerative
retinas [Homo sapiens] (SEQ ID NO:90)
MTPWLGLIVLLGSWSLGDWGAEACTCSPSHPQDAFCNSDIVIRAKVVGKKLVKEGPFGTLVYTIKQMKMY
RGFTKMPHVQYIHTEASESLCGLKLEVNKYQYLLTGRVYDGKMYTGLCNFVERWDQLTLSQRKGLNYRYH
LGCNCKIKSCYYLPCFVTSKNECLWTDMLSNFGYPGYQSKHYACIRQKGGYCSWYRGWAPPDKSIINATD
P

>gi|21536431|ref|NM_000362.3| Homo sapiens tissue inhibitor of
metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory) (TIMP3),
mRNA (SEQ ID NO:91)
```
TCTGTCGACTTGCCCCAGAGCTGATCCTTGTCTTTGTCCACTTCTCAGCGAGGATGGCACTTCAGGGAGC
CCTTCCCTTACTATCGCAGAGAGAGCAGGCCCTCCCCAGTCATGTCCAACCCAGAACTCTGTTTTGTTTT
CTTCATAGCCCTAGCATCACAGAAAATCACCCTGTGCATTCATGGATGTCCACGGGGGCAAGGGCTTTGT
GTTGCTTAACCCAGCATCCTGAACCGTGTTTGTTGAATGAATACAGAACCCCGTTTGCTCTGGGAGAGCA
CAGAAAACAGTCTTCTATCATATCATAGCCAGCTGCAAACAGCAGATGGCTTCCCATATCCCAGAGAG
TAAGAACCAGAGAGAGAGAGAAAGAGAGAGTTTGGGTCTTTCTCCTCTGTCCTGCTCTCTCCAGAGA
AACTGGAGGGGTAGCAGTTAGCATTCCCCCGCTGGTTCCACCAAGCACAGTCAAGGTCTCTAGGACATGG
CCACCCCTCACCTGTGGAAGCGGTCCTGCTGGGGTGGGTGGGTGTTAGTTGGTTCTGGTTTGGGTCAGAG
ACACCCAGTGGCCCAGGTGGGCGTGGGGCCAGGGCGCAGACGAGAAGGGGCACGAGGGCTCCGCTCCGAG
GACCCAGCGGCAAGCACCGGTCCCGGGCGCGCCCCAGCCCACCCACTCGCGTGCCCACGGCGGCATTATT
CCCTATAAGGATCTGAACGATCCGGGGGCGGCCCCGCCCCGTTACCCCTTGCCCCCGGCCCCGCCCCCTT
TTTGGAGGGCCGATGAGGTAATGCGGCTCTGCCATTGGTCTGAGGGGGCGGGCCCCAACAGCCCGAGGCG
GGGTCCCCGGGGGCCCAGCGCTATATCACTCGGCCGCCCAGGCAGCGGCGCAGAGCGGGCAGCAGGCAGG
CGGCGGGCGCTCAGACGGCTTCTCCTCCTCCTCTTGCTCCTCCAGCTCCTGCTCCTTCGCCGGGAGGCCG
CCCGCCGAGTCCTGCGCCAGCGCCGAGGCAGCCTCGCTGCGCCCCATCCCGTCCGCCGGGCACTCGGAG
GGCAGCGCGCCGGAGGCCAAGGTTGCCCCGCACGGCCCGGCGGGCGAGCGAGCTCGGGCTGCAGCAGCCC
CGCCGGCGGCGCGCACGGCAACTTTGGAGAGGCGAGCAGCAGCCCCGGCAGCGGCGGCAGCAGCGGCAAT
GACCCCTTGGCTCGGGCTCATCGTGCTCCTGGGCAGCTGGAGCCTGGGGGACTGGGGCGCCGAGGCGTGC
ACATGCTCGCCCAGCCACCCCCAGGACGCCTTCTGCAACTCCGACATCGTGATCCGGGCCAAGGTGGTGG
GGAAGAAGCTGGTAAAGGAGGGGCCCTTCGGCACGCTGGTCTACACCATCAAGCAGATGAAGATGTACCG
```

```
AGGCTTCACCAAGATGCCCCATGTGCAGTACATCCACACGGAAGCTTCCGAGAGTCTCTGTGGCCTTAAG
CTGGAGGTCAACAAGTACCAGTACCTGCTGACAGGTCGCGTCTATGATGGCAAGATGTACACGGGGCTGT
GCAACTTCGTGGAGAGGTGGGACCAGCTCACCCTCTCCCAGCGCAAGGGGCTGAACTATCGGTATCACCT
GGGTTGTAACTGCAAGATCAAGTCCTGCTACTACCTGCCTTGCTTTGTGACTTCCAAGAACGAGTGTCTC
TGGACCGACATGCTCTCCAATTTCGGTTACCCTGGCTACCAGTCCAAACACTACGCCTGCATCCGGCAGA
AGGGCGGCTACTGCAGCTGGTACCGAGGATGGGCCCCCCCGGATAAAAGCATCATCAATGCCACAGACCC
CTGAGCGCCAGACCCTGCCCCACCTCACTTCCCTCCCTTCCCGCTGAGCTTCCCTTGGACACTAACTCTT
CCCAGATGATGACAATGAAATTAGTGCCTGTTTTCTTGCAAATTTAGCACTTGGAACATTTAAAGAAAGG
TCTATGCTGTCATATGGGGTTTATTGGGAACTATCCTCCTGGCCCCACCCTGCCCCTTCTTTTGGTTTT
GACATCATTCATTTCCACCTGGGAATTTCTGGTGCCATGCCAGAAAGAATGAGGAACCTGTATTCCTCTT
CTTCGTGATAATATAATCTCTATTTTTTTAGGAAAACAAAATGAAAAACTACTCCATTTGAGGATTGTA
ATTCCCACCCCTCTTGCTTCTTCCCCACCTCACCATCTCCCAGACCCTCTTCCCTTTGCCCTTCTCCTCC
AATACATAAAGGACACAGACAAGGAACTTGCTGAAAGGCCAACCATTTCAGGATCAGTCAAAGGCAGCAA
GCAGATAGACTCAAGGTGTGTGAAAGATGTTATACACCAGGAGCTGCCACTGCATGTCCCAACCAGACTG
TGTCTGTCTGTGTCTGCATGTAAGAGTGAGGGAGGGAAGGAAGGAACTACAAGAGAGTCGGAGATGATGC
AGCACACACACAATTCCCCAGCCCAGTGATGCTTGTGTTGACCAGATGTTCCTGAGTCTGGAGCAAGCAC
CCAGGCCAGAATAACAGAGCTTTCTTAGTTGGTGAAGACTTAAACATCTGCCTGAGGTCAGGAGGCAATT
TGCCTGCCTTGTACAAAAGCTCAGGTGAAAGACTGAGATGAATGTCTTTCCTCTCCCTGCCTCCCACCAG
ACTTCCTCCTGGAAAACGCTTTGGTAGATTTGGCCAGGAGCTTTCTTTTATGTAAATTGGATAAATACAC
ACACCATACACTATCCACAGATATAGCCAAGTAGATTTGGGTAGAGGATACTATTTCCAGAATAGTGTTT
AGCTCACCTAGGGGGATATGTTTGTATACACATTTGCATATACCCACATGGGGACATAAGCTAATTTTTT
TACAGGACACAGAATTCTGTTCAATGCTGTTAAATATGCCAATAGTTTAATCTCTTCTATTTTGTTGTCG
TTGCTTGTTTGAAGAAAATCATGACATTCCAAGTTGACATTTTTTTTTCATTTTAATTAAAATTTGAAAT
TCTGAACACCGTCAGCACCCTCTCTTCCCTATCATGGGTCATCTGACCCCTGTCCGTCTCCTTGTCCCTG
CTTCATGTTTGGGGGCCTTTCTTTAACTGCCTTCCTGGCTTAGCTCAGATGGCAGATGAGAGTGTAGTCA
AGGGCCTGGGCACAGGAGGGAGAGCTGCAGAGTGTCCTGCCTGCCTTGGCTGGAGGGACACCTCTCCTGG
GTGTGGAGACAGCTTGGTTCCCTTTCCCTAGCTCCCTGGTGGGTGAATGCCACCTCCTGAGATCCTCACC
TCTTGGAATTAAAATTGTTGGTCACTGGGGAAAGCCTGAGTTTGCAACCAGTTGTAGGGTTTCTGTTGTG
TTTTTTTTTTTTTTTGAAATAAAACTATAATATAAATTCTCCTATTAAATAAAATTATTTTAAGTTTT
AGTGTCAAAAGTGAGATGCTGAGAGTAGGTGATAATGTATATTTTACAGAGTGGGGGTTGGCAGGATGGT
GACATTGAACATGATTGCTCTCTGTCTCTTTTTTCAGCTTATGGGTATTTATCTTCTATTAGTATTTGTA
TCTTCAGTTCATTCCACTTTAGGAAACAGAGCTGCCAATTGAAACAGAAGAAGAAAAAAAAAAAAAAGCAG
CAGACAACACACTGTAGAGTCTTGCACACACACAAGTGCCCAGGCAAGGTGCTTGGCAGAACCGCAGAGT
GGGAAGAGAGTACCGGCATCGGGTTTCCTTGGGATCAATTTCATTACCGTGTACCTTTCCCATTGTGGTC
ATGCCATTTGGCAGGGGGAGAATGGGAGGCTTGGCCTTCTTTGTGAGGCAGTGTGAGCAGAAGCTGATGC
CAGCATGTCACTGGTTTTGAAGGGATGAGCCCAGACTTGATGTTTTGGGATTGTCCTTATTTTAACCTCA
AGGTCTCGCATGGTGGGGCCCCTGACCAACCTACACAAGTTCCCTCCCACAAGTGGACATCAGTGTCTTC
TCTGTGAGGCATCTGGCCATTCGCACTCCCTGGTGTGGTCAGCCTCTCTCACACAAGGAGGAACTTGGGT
GAAGGCTGAGTGTGAGGCACCTGAAGTTTCCCTGCGGAGTCGATAAATTAGCAGAACCACATCCCCATCT
GTTAGGCCTTGGTGAGGAGGCCCTGGGCAAAGAAGGGTCTTTCGCAAAGCGATGTCAGAGGGCGGTTTTG
AGCTTTCTATAAGCTATAGCTTTGTTTATTTCACCCGTTCACTTACTGTATAATTTAAAATCATTTATGT
AGCTGAGACACTTCTGTATTTCAATCATATCATGAACATTTTATTTTGCTAAATCTTGTGTCATGTGTAG
GCTGTAATATGTGTACATTGTGTTTAAGAGAAAAATGAAACCCACATGCCGCCATTTTCCTGAATCAAAT
TCTGCAGTGGAATGGAGAGGAAAATACTTCTAGGCAAGCAGCTAGACTGGTGAATTGGGGGAAATAGAAG
GAACTAGTAACTGAGACTCCTCCAGCCTCCTCCCTATTGGAATCCCAATGGCTCCTGGAGTAGGAAAAAA
GTTTAAACTACATTCATGTTCTTGTTCTGTGTCACTCGGCCCTGGGTAGTCTACCATTTACTTCACCCCA
AGTCCTGCTGCCCATCCAGTTGGGAAGCCATGATTTTCCTAAGAATCCAGGGCCATGGGAGATACAATTC
CAAGTTCTCGCTTCCTCCTTTGGGCATCTCTTCTGCCTCCCAATCAAGGAAGCTCCATGCTCAGGCTCTC
AGCTCTCGGGCCAGTGCTCTGCTCTGTCCAGGGTAGGTAATACTGGGAGACTCCTGTCTTTTACCCTCCC
CTCGTTCCAGACCTGCCTCATGGTGGCAACATGGTTCTTGAACAATTAAAGAAACAAATGACTTTTTGGA
ATAGCCCTGTCTAGGGCAAACTGTGGCCCCCAGGAGACACTACCCTTCCATGCCCCAGACCTCTGTCTTG
CATGTGACAATTGACAATCTGGACTACCCCAAGATGGCACCCAAGTGTTTGGCTTCTGGCTACCTAAGGT
TAACATGTCACTAGAGTATTTTTATGAGAGACAAACATTATAAAAATCTGATGGCAAAAGCAAACAAAA
TGGAAAGTAGGGGAGGTGGATGTGACAACAACTTCCAAATTGGCTCTTTGGAGGCGAGAGGAAGGGGAGA
ACTTGGAGAATAGTTTTTGCTTTGGGGGTAGAGGCTTCTTAGATTCTCCCAGCATCCGCCTTTCCCTTTA
GCCAGTCTGCTGTCCTGAAACCCAGAAGTGATGGAGAGAAACCAACAAGAGATCTCGAACCCTGTCTAGA
AGGAATGTATTTGTTGCTAAATTTCGTAGCACTGTTTACAGTTTTCCTCCATGTTATTTATGAATTTTAT
ATTCCGTGAATGTATATTGTCTTGTAATGTTGCATAATGTTCACTTTTTATAGTGTGTCCTTTATTCTAA
ACAGTAAAGTGGTTTTATTTCTATCAC
```

RNH

>gi|4506564|ref|NM_002939.1| Homo sapiens ribonuclease/angiogenin inhibitor
(RNH), mRNA (SEQ ID NO:92)
AGTGCCTTAGATTCCAGCGAGCTACGAAGCAATCCTGGCCCAGCCGAGCTTGCTTCCCCAAATCCCGTAA
TCCTTGACCTTATTCCCCCAAAGAAGCGGCCTCCCGGGAAGGAGCGCCCTGGCGGAGAAGACTCGAACGG
CTCCCACAGCCGGGCGTTGGGGGTAAAGGCATGAAGAACTCTTGACTGACAGAAACGGAGGGTGTGTCCA
AAGTTTTGAGGACGGCCGAGCGGCGCTCCAAAACCCGTCCTCACAGCCTCGCCCCGTTCGCCTCAGCTAC
AACAAATCATCGTCAACCTGTTCCACCTTCTCCAGTCTGGTAGCAAAAAGGGGTGTCTCAGGCCACTCTT
CACCTCCACC<u>ATG</u>AGCCTGGACATCCAGAGCCTGGACATCCAGTGTGAGGAGCTGAGCGACGCTAGATGG
GCCGAGCTCCTCCCTCTGCTCCAGCAGTGCCAAGTGGTCAGGCTGGACGACTGTGGCCTCACGGAAGCAC
GGTGCAAGGACATCAGCTCTGCACTTCGAGTCAACCCTGCACTGGCAGAGCTCAACCTGCGCAGCAACGA
GCTGGGCGATGTCGGCGTGCATTGCGTGCTCCAGGGCCTGCAGACCCCCTCCTGCAAGATCCAGAAGCTG
AGCCTCCAGAACTGCTGCCTGACGGGGGCCGGCTGCGGGGTCCTGTCCAGCACACTACGCACCCTGCCCA
CCCTGCAGGAGCTGCACCTCAGCGACAACCTCTTGGGGGATGCGGGCCTGCAGCTGCTCTGCGAAGGACT
CCTGGACCCCCAGTGCCGCCTGGAAAAGCTGCAGCTGGAGTATTGCAGCCTCTCGGCTGCCAGCTGCGAG
CCCCTGGCCTCCGTGCTCAGGGCCAAGCCGGACTTCAAGGAGCTCACGGTTAGCAACAACGACATCAATG
AGGCTGGCGTCCGTGTGCTGTGCCAGGGCCTGAAGGACTCCCCCTGCCAGCTGGAGGCGCTCAAGCTGGA
GAGCTGCGGTGTGACATCAGACAACTGCCGGGACCTGTGCGGCATTGTGGCCTCCAAGGCCTCGCTGCGG
GAGCTGGCCCTGGGCAGCAACAAGCTGGGTGATGTGGGCATGGCGGAGCTGTGCCCAGGGCTGCTCCACC
CCAGCTCCAGGCTCAGGACCCTGTGGATCTGGGAGTGTGGCATCACTGCCAAGGGCTGCGGGGATCTGTG
CCGTGTCCTCAGGGCCAAGGAGAGCCTGAAGGAGCTCAGCCTGGCCGGCAACGAGCTGGGGGATGAGGGT
GCCCGACTGCTGTGTGAGACCCTGCTGGAACCTGGCTGCCAGCTGGAGTCGCTGTGGGTGAAGTCCTGCA
GCTTCACAGCCGCCTGCTGCTCCCACTTCAGCTCAGTGCTGGCCCAGAACAGGTTTCTCCTGGAGCTACA
GATAAGCAACAACAGGCTGGAGGATGCGGGCGTGCGGGAGCTGTGCCAGGGCCTGGGCCAGCCTGGCTCT
GTGCTGCGGGTGCTCTGGTTGGCCGACTGCGATGTGAGTGACAGCAGCTGCAGCAGCCTCGCCGCAACCC
TGTTGGCCAACCACAGCCTGCGTGAGCTGGACCTCAGCAACAACTGCCTGGGGGACGCGGGCATCCTGCA
GCTGGTGGAGAGCGTATCCGAGCCGGGCTGCCTCCTGGAGCAGCTGGTCCTGTACGACATTTACTGGTCT
GAGGAGATGGAGGACCGGCTGCAGGCCCTGGAGAAGGACAAGCCATCCCTGAGGGTCATCTCC<u>TGA</u>AGCT
CTTCCTGCTGCTGCTCTCCCTGGACGACCGGCCTCGAGGCAACCCTGGGGCCCACCAGCCCCTGCCATGC
TCTCACCCTGCATATCCTAGGTTTGAAGAGAAACGCTCAGATCCGCTTATTTCTGCCAGTATATTTTGGA
CACTTTATAATCATTAAAGCACTTTCTTGGC

>gi|4506565|ref|NP_002930.1| ribonuclease/angiogenin inhibitor; Placental
ribonuclease inhibitor [Homo sapiens] (SEQ ID NO:93)
MSLDIQSLDIQCEELSDARWAELLPLLQQCQVVRLDDCGLTEARCKDISSALRVNPALAELNLRSNELGD
VGVHCVLQGLQTPSCKIQKLSLQNCCLTGAGCGVLSSTLRTLPTLQELHLSDNLLGDAGLQLLCEGLLDP
QCRLEKLQLEYCSLSAASCEPLASVLRAKPDFKELTVSNNDINEAGVRVLCQGLKDSPCQLEALKLESCG
VTSDNCRDLCGIVASKASLRELALGSNKLGDVGMAELCPGLLHPSSRLRTLWIWECGITAKGCGDLCRVL
RAKESLKELSLAGNELGDEGARLLCETLLEPGCQLESLWVKSCSFTAACCSHFSSVLAQNRFLLELQISN
NRLEDAGVRELCQGLGQPGSVLRVLWLADCDVSDSSCSSLAATLLANHSLRELDLSNNCLGDAGILQLVE
SVS<u>E</u>PGCLLEQLVLYDIYWSEEMEDRLQALEKDKPSLRVIS

>gi|21361546|ref|NM_002939.2| Homo sapiens ribonuclease/angiogenin
inhibitor (RNH), mRNA (SEQ ID NO:94)
CACACCCTGGTCGGTTTTGCTCACGTGCTTCGGGTCGGTTGGATTCAGTTCCTCCATGTGGGGCCGTGGG
ATGTCACCACCCTTTGCCACTGTCCTTGCAGGCCGGGCGCCCAGGCGTGTGGACCCGAGCCCAGCCTTGC
GTCTTTCCTCCCAGTTGGTACCAGGGGCCTTGTGCCTCCACTCATGCAGAGTGGACCAGCCGCCTCTGAA
GCAGCCCCGGGGAGAGGAGTGCCAGGCACACAGACTCAGGAAGCTGGGGGTGTCAGGGCACCAGGAGCCT
CTCAAGCGGTCCTGTTGGCCGTCTTTCTGAAGGGCAGCTGGGGTTGCATCATTTCCCCAAGTGGATCCTG
CCAACTTTCTGGGGCTTCTGGATGAGAGCCCTCTCCCCACCCCACCCACCCCTCCCCGCATTGTGGCATC
AGTGCTGCTGCTTCCAGGGAGCCTTCCTGGCCATCCAAGCCTCCTCTCCAGGTTCCTGCCCTGCTGTAGT
CCCCAGGCCAGTGCTTGGCAGGTGCTCAGGGAATGTATCCACCAACCAAGGTTTGGGGTGGCTGTCTCTG
CCTGACCACTTTCCCCAGGCCCCTGGCGGGTACCTGAGCTGTGCTCTCAGGGCCTCAGGAACCTCCTTCC
ATATTCAGGGCCTGTGCCTGGGGAGGCTTCAGGGTGTAGCAGCTGTGCCCATCCCAGGCTGACCCACCCA
GCTTGCCTGGTAGCCCAGCCTCTGGGCTAGTGTGCCGTGGGGCAGGGGATGTGCTGTAGCCTGGTGCAGA
GTCCCCAACCCCAGAAGGGGGCCATGGAAGCTGACACCCCAAGTGGCCGCCCCCCTGCTCTGTCTTGCTT
CGGACACTGTGGCCGGGTCCAGGATCCTGGCATCCTGGGAGGTCTCTGGCTTTGTGGGCAGCCTGCCTGG
CCCGCACAGTCTGCCTGTCCTGAGGGTGAGACACAGGTCAAGCCCACAGACCCCCTTGCTCCCCTGCTGG
GGCCTCCAGGCTCACAGACCACCCCACCCTACCCTGTCCTTGCCCAAGCAAATG<u>A</u>GAGGCAGGGGCTTCC
CGGGCTGCTGCTGTCCCGCCCTCTGTGGGGCAGGAGGAGGTGCCCACAGAGGCTGGGTGGTGATAGCCAG

GAGATGGGCTGGCATCTGCATTACCCAAGCTCTGCTGCCCATGGTGGCCTTTCTGGGGGTGGGTGCTGGT
CCCTGCCCCCTGCCCCACCCCTGATGTCTGCTCCAGAGACAAAGGTGGGGAGGGTGCTGAAGAGGAAGTG
TTTGCCCAGGGAGAGGCTGCGGCTCCTCCTGAAACATCAGCCCTGTGGGTCCTGTTTGCAGAATCTCCGG
CCTGTGAAACTGTGAGGGGATTCAGCCAAGACGTCCTCTTCCCTCTGCCTCCCACCCAGGCCACTCTTCA
CCTCCACC<u>ATG</u>AGCCTGGACATCCAGAGCCTGGACATCCAGTGTGAGGAGCTGAGCGACGCTAGATGGGC
CGAGCTCCTCCCTCTGCTCCAGCAGTGCCAAGTGGTCAGGCTGGACGACTGTGGCCTCACGGAAGCACGG
TGCAAGGACATCAGCTCTGCACTTCGAGTCAACCCTGCACTGGCAGAGCTCAACCTGCGCAGCAACGAGC
TGGGCGATGTCGGCGTGCATTGCGTGCTCCAGGGCCTGCAGACCCCCTCCTGCAAGATCCAGAAGCTGAG
CCTCCAGAACTGCTGCCTGACGGGGGCCGGCTGCGGGTCCTGTCCAGCACACTACGCACCCTGCCCACC
CTGCAGGAGCTGCACCTCAGCGACAACCTCTTGGGGGATGCGGGCCTGCAGCTGCTCTGCGAAGGACTCC
TGGACCCCCAGTGCCGCCTGGAAAAGCTGCAGCTGGAGTATTGCAGCCTCTCGGCTGCCAGCTGCGAGCC
CCTGGCCTCCGTGCTCAGGGCCAAGCCGGACTTCAAGGAGCTCACGGTTAGCAACAACGACATCAATGAG
GCTGGCGTCCGTGTGCTATGCCAGGGCCTGAAGGACTCCCCCTGCCAGCTGGAGGCGCTCAAGCTGGAGA
GCTGCGGTGTGACATCAGACAACTGCCGGGACCTGTGCGGCATTGTGGCCTCCAAGGCCTCGCTGCGGGA
GCTGGCCCTGGGCAGCAACAAGCTGGGTGATGTGGGCATGGCGGAGCTGTGCCCAGGGCTGCTCCACCCC
AGCTCCAGGCTCAGGACCCTGTGGATCTGGGAGTGTGGCATCACTGCCAAGGGCTGCGGGGATCTGTGCC
GTGTCCTCAGGGCCAAGGAGAGCCTGAAGGAGCTCAGCCTGGCCGGCAACGAGCTGGGGGATGAGGGTGC
CCGACTGCTGTGTGAGACCCTGCTGGAACCTGGCTGCCAGCTGGAGTCGCTGTGGGTGAAGTCCTGCAGC
TTCACAGCCGCCTGCTGCTCCCACTTCAGCTCAGTGCTGGCCCAGAACAGGTTTCTCCTGGAGCTACAGA
TAAGCAACAACAGGCTGGAGGATGCGGGCGTGCGGGAGCTGTGCCAGGGCCTGGGCCAGCCTGGCTCTGT
GCTGCGGGTGCTCTGGTTGGCCGACTGCGATGTGAGTGACAGCAGCTGCAGCAGCCTCGCCGCAACCCTG
TTGGCCAACCACAGCCTGCGTGAGCTGGACCTCAGCAACAACTGCCTGGGGGACGCGGGCATCCTGCAGC
TGGTGGAGAGCGTCCGGCAGCCGGGCTGCCTCCTGGAGCAGCTGGTCCTGTACGACATTTACTGGTCTGA
GGAGATGGAGGACCGGCTGCAGGCCCTGGAGAAGGACAAGCCATCCCTGAGGGTCATCTCC<u>TGA</u>GGCTCT
TCCTGCTGCTGCTCTCCCTGGACGACCGGCCTCGAGGCAACCCTGGGGCCCACCAGCCCCTGCCATGCTC
TCACCCTGCATATCCTAGGTTTGAAGAGAAACGCTCAGATCCGCTTATTTCTGCCAGTATATTTTGGACA
CTTTATAATCATTAAAGCACTTTCTTGGCAGGAAAAAAAAA

>gi|21361547|ref|NP_002930.2| ribonuclease/angiogenin inhibitor; Placental
ribonuclease inhibitor [Homo sapiens] (SEQ ID NO:95)
MSLDIQSLDIQCEELSDARWAELLPLLQQCQVVRLDDCGLTEARCKDISSALRVNPALAELNLRSNELGD
VGVHCVLQGLQTPSCKIQKLSLQNCCLTGAGCGVLSSTLRTLPTLQELHLSDNLLGDAGLQLLCEGLLDP
QCRLEKLQLEYCSLSAASCEPLASVLRAKPDFKELTVSNNDINEAGVRVLCQGLKDSPCQLEALKLESCG
VTSDNCRDLCGIVASKASLRELALGSNKLGDVGMAELCPGLLHPSSRLRTLWIWECGITAKGCGDLCRVL
RAKESLKELSLAGNELGDEGARLLCETLLEPGCQLESLWVKSCSFTAACCSHFSSVLAQNRFLLELQISN
NRLEDAGVRELCQGLGQPGSVLRVLWLADCDVSDSSCSSLAATLLANHSLRELDLSNNCLGDAGILQLVE
SVRQPGCLLEQLVLYDIYWSEEMEDRLQALEKDKPSLRVIS

WHAT IS CLAIMED IS:

1   1.   A method for identifying a compound that modulates T

2   lymphocyte activation, the method comprising the steps of:

3   (i) contacting the compound with an A-raf-1 or TCPTP/PTPN2

4   polypeptide or a fragment thereof, the polypeptide or fragment thereof encoded by a

5   nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding a

6   polypeptide having an amino acid sequence of SEQ ID NO:2, 4, or 28; and

7   (ii) determining the functional effect of the compound upon the A-raf-1 or

8   TCPTP/PTPN2 polypeptide.

1   2.   The method of claim 1, wherein the functional effect is measured

2   *in vitro*.

1   3.   The method of claim 2, wherein the functional effect is a physical

2   effect.

1   4.   The method of claim 3, wherein the functional effect is determined

2   by measuring ligand or substrate binding to the polypeptide.

1   5.   The method of claim 2, wherein the functional effect is a chemical

2   effect.

1   6.   The method of claim 1, wherein the polypeptide is expressed in a

2   host cell.

1   7.   The method of claim 6, wherein the functional effect is a physical

2   effect.

1   8.   The method of claim 7, wherein the functional effect is determined

2   by measuring ligand or substrate binding to the polypeptide.

1   9.   The method of claim 6, wherein the functional effect is a chemical

2   or phenotypic effect.

1   10.   The method of claim 6, wherein the host cell is primary T

2   lymphocyte.

1   11.   The method of claim 6,wherein the host cell is a cultured T cell.

1   12.   The method of claim 11, wherein the host cell is a Jurkat cell.

1   13.   The method of claim 6, wherein the chemical or phenotypic effect

2   is determined by measuring CD69 expression, intracellular Ca^{2+} mobilization, Ca^{2+}

3   influx, or lymphocyte proliferation.

1   14.   The method of claim 1, wherein modulation is inhibition of T

2   lymphocyte activation.

1   15.   The method of claim 1, wherein the polypeptide is recombinant.

1   16.   The method of claim 1, wherein the compound is an antibody.

1   17.   The method of claim 1, wherein the compound is an antisense

2   molecule.

1   18.   The method of claim 1, wherein the compound is a RNAi

2   molecule.

1   19.   The method of claim 1, wherein the compound is a small organic

2   molecule.

1   20.   The method of claim 1, wherein the compound is a peptide.

1   21.   The method of claim 20, wherein the peptide is circular.

1   22.   A method for identifying a compound that modulates T

2   lymphocyte activation, the method comprising the steps of:

3   (i) contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2

4   polypeptide or fragment thereof with the compound, the A-raf-1 or TCPTP/PTPN2

5   polypeptide or fragment thereof encoded by a nucleic acid that hybridizes under stringent

6   conditions to a nucleic acid encoding a polypeptide having an amino acid sequence of

7   SEQ ID NO:2, 4, or 28; and

8   (ii) determining the chemical or phenotypic effect of the compound upon

9   the cell comprising the A-raf-1 or TCPTP/PTPN2 polypeptide or fragment thereof,

10   thereby identifying a compound that modulates T lymphocyte activation.

1   23.   A method for identifying a compound that modulates T
2   lymphocyte activation, the method comprising the steps of:
3   (i) contacting the compound with an A-raf-1 or TCPTP/PTPN2
4   polypeptide or a fragment thereof, the A-raf-1 or TCPTP/PTPN2 polypeptide or fragment
5   thereof encoded by a nucleic acid that hybridizes under stringent conditions to a nucleic
6   acid encoding a polypeptide having an amino acid sequence of SEQ ID NO:2, 4, or 28;
7   (ii) determining the physical effect of the compound upon the A-raf-1 or
8   TCPTP/PTPN2 polypeptide; and
9   (iii) determining the chemical or phenotypic effect of the compound upon
10   a cell comprising the A-raf-1 or TCPTP/PTPN2 polypeptide or fragment thereof, thereby
11   identifying a compound that modulates T lymphocyte activation.

1   24.   A method of modulating T lymphocyte activation in a subject, the
2   method comprising the step of administering to the subject a therapeutically effective
3   amount of a compound identified using the method of claim 1.

1   25.   The method of claim 24, wherein the subject is a human.

1   26.   The method of claim 24, wherein the compound is an antibody.

1   27.   The method of claim 24, wherein the compound is an antisense
2   molecule.

1   28.   The method of claim 24, wherein the compound is a RNAi
2   molecule.

1   29.   The method of claim 24, wherein the compound is a small organic
2   molecule.

1   30.   The method of claim 24, wherein the compound is a peptide.

1   31.   The method of claim 30, wherein the peptide is circular.

1   32.   The method of claim 24, wherein the compound inhibits T
2   lymphocyte activation.

1   33.   A method of modulating T lymphocyte activation in a subject, the
2   method comprising the step of administering to the subject a therapeutically effective
3   amount of an A-raf-1 or TCPTP/PTPN2 polypeptide, the polypeptide encoded by a
4   nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding a
5   polypeptide having an amino acid sequence of SEQ ID NO:2, 4, or 28.

1   34.   A method of modulating T lymphocyte activation in a subject, the
2   method comprising the step of administering to the subject a therapeutically effective
3   amount of a nucleic acid encoding an A-raf-1 or TCPTP/PTPN2 polypeptide, wherein the
4   nucleic acid hybridizes under stringent conditions to a nucleic acid encoding a
5   polypeptide having an amino acid sequence of SEQ ID NO:2, 4, or 28.

Identification of Regulatory Proteins
that Affect T-Cell Activation
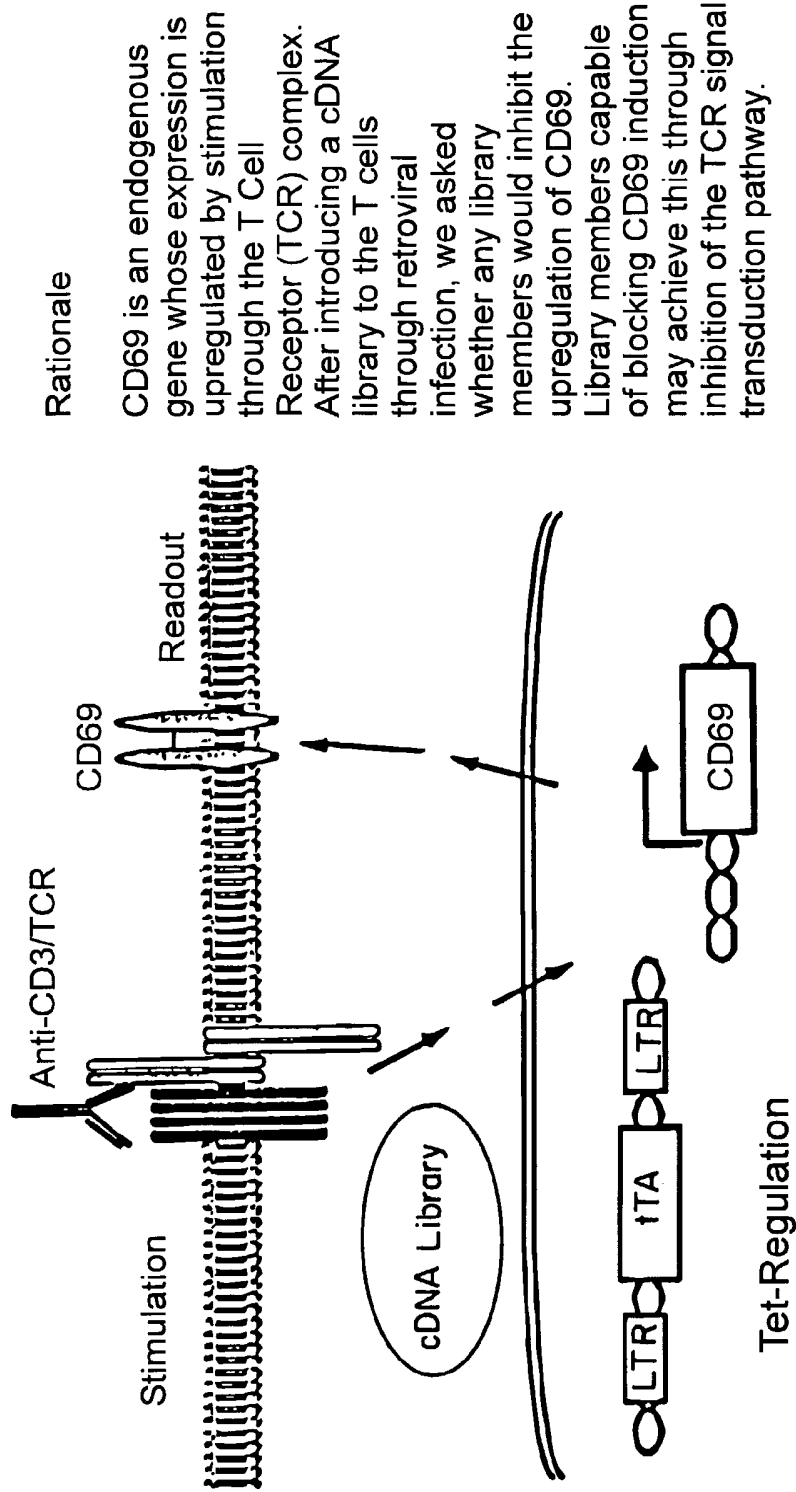
Rationale

CD69 is an endogenous gene whose expression is upregulated by stimulation through the T Cell Receptor (TCR) complex. After introducing a cDNA library to the T cells through retroviral infection, we asked whether any library members would inhibit the upregulation of CD69. Library members capable of blocking CD69 induction may achieve this through inhibition of the TCR signal transduction pathway.

FIG. 1

Surface Marker Screen
TCR Activation-Induced Expression of CD69

Infect Jurkat N (Tet-regulatable line)

↓

Stimulate cells with anti-TCR mAb

↓

FACS sort for CD69^{low}CD3$^+$ cells

↓

Recover and expand desired cells

Repeat until phenotype is established

↓

Single cell cloning

↓

Tet-regulatable phenotype

↓

Rescue DNA
Sequence
Transfer to naive cells

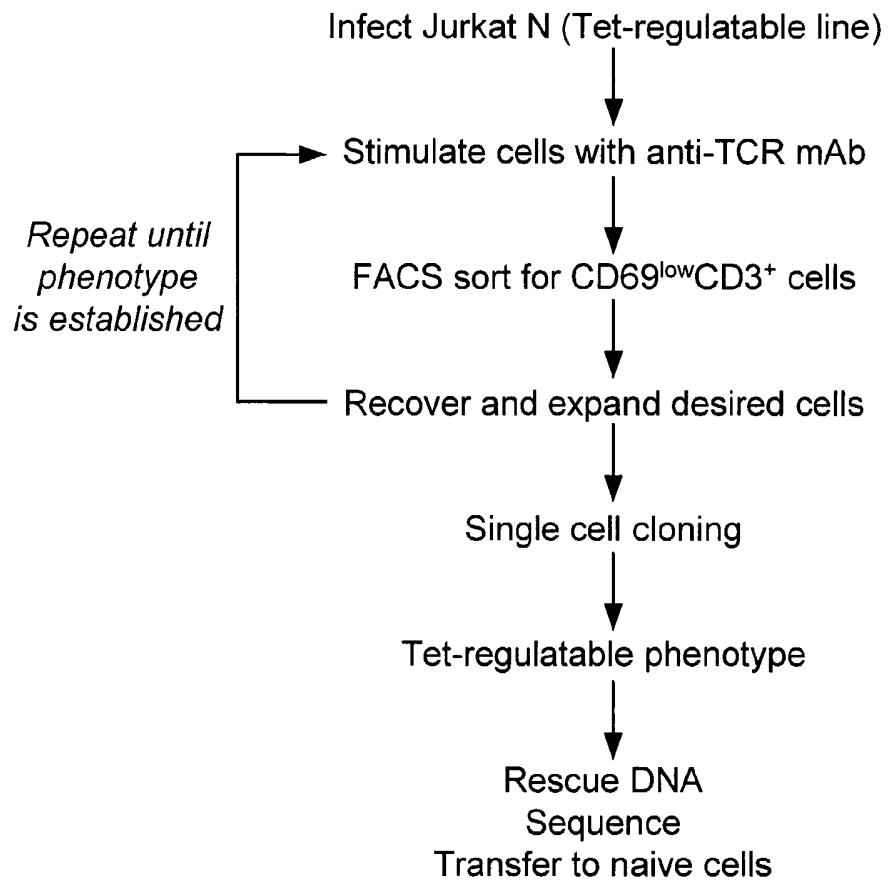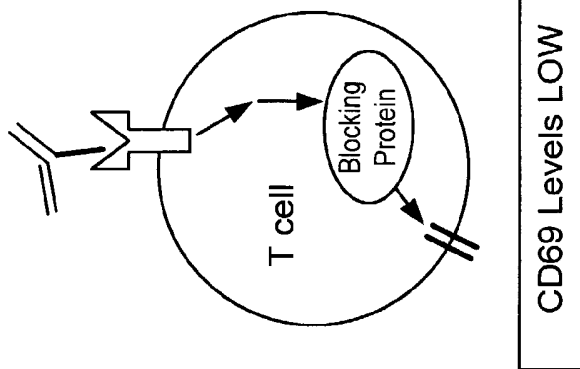**FIG. 2**

Distinguish cDNA-induced Phenotype from Somatic Mutations

With Doxycycline (+Dox)
Library cDNA Expression Off

T cell

CD69 Levels HIGH

Stimulate T cells

Compare CD69 expression

Without Doxycycline (-Dox)
Library cDNA Expression On
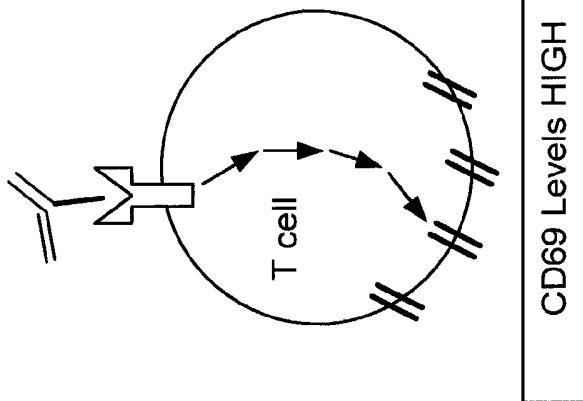
T cell

Blocking Protein

CD69 Levels LOW

FIG. 3

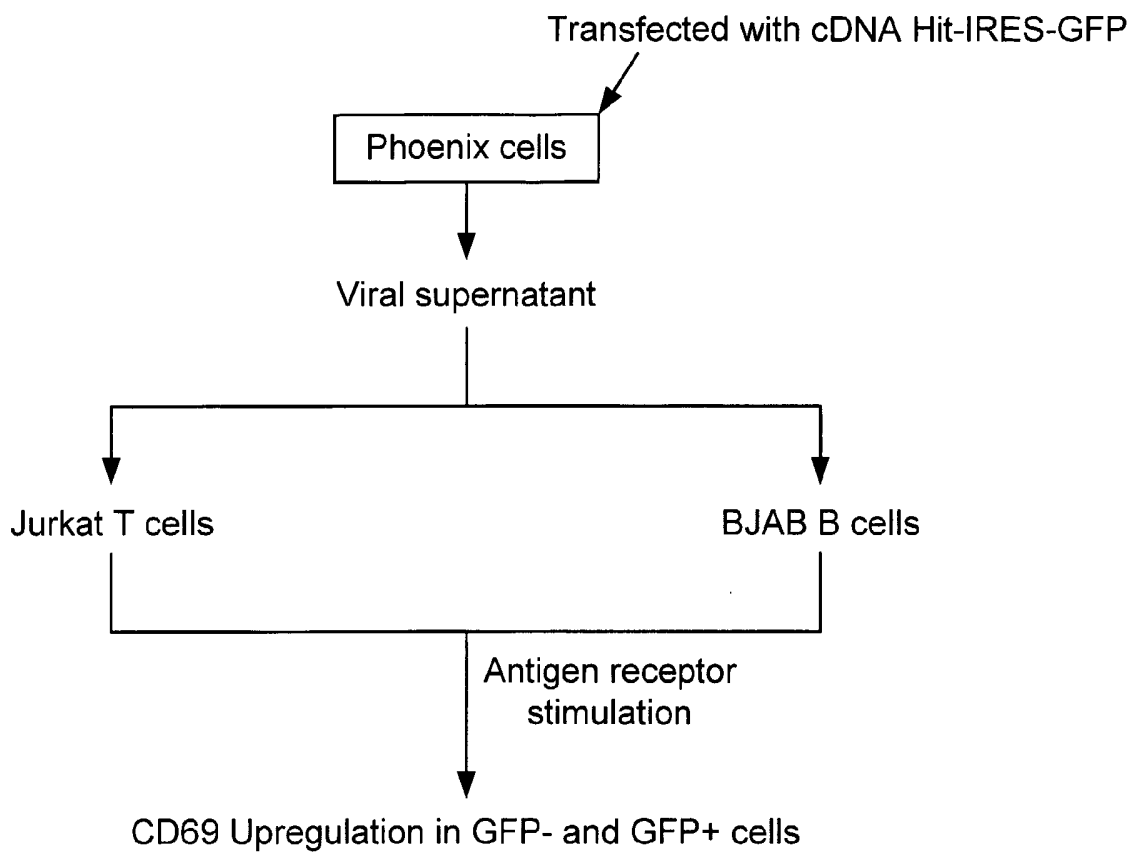## Cell Specificity of the cDNA Hits

FIG. 4

Known TCR Regulators Uncovered from CD69 cDNA Screen c3.1

Identify	Classification	Protein Family	Successful Phenotype Transfer of DNA	Direction	Length
TCRβ (multiple)	Known	receptor	on-hold	Sense	Partial
LCK	Known	Tyr kinase	Yes	Sense	Partial
ZAP70	Known	Tyr kinase	Yes	Sense	Partial
SYK	Known	Tyr kinase	Yes	Sense	Partial
PLCγ1	Known	phospholipase	Yes	Sense	Partial
PAG/cbp	Known	TM adaptor	on-hold	Sense	Partial
A-raf-1	Known	Ser/Thr kinase	on-hold	Sense	Partial
SHP/PTP1C	Known	Ser/Thr phosphatase	on-hold	Sense	Partial
CSK	Known	Tyr kinase	on-hold	Sense	Partial
Nucleolin	Known	RNA-binding	on-hold	Sense	Partial

FIG. 5

6/8

Primary Potential Hits from CD69 cDNA Screen c3.1

Identify	Protein Family	Successful Phenotype Transfer	Direction	Length
Novel 1	Ser/Thr kinase	Yes	Sense	Partial
Novel 2	Tyr phosphatase	Yes	Sense	Partial
Novel 3	GPCR (receptor)	Yes	Sense	Partial
Novel 4	cytokine receptor	Yes	Sense	Partial
Novel 5	trans-membrane receptor	Yes	Sense	Partial
Novel 6	TNF-α induced mRNA	Yes	Sense	full-length
Novel 7	Zinc finger	Yes	Sense	Partial

FIG. 6

7/8

Overview of identified molecular targets

gene	domain homology	direction	accession #	Length	frequency
known to TCR pathway					
TCRb	receptor		numerous	partial	48
Lck	Tyr kinase		U23852 1	nt 1-859	4
ZAP70	Tyr kinase		L05148 1	nt 63-997	12
ZAP70(long)	Tyr kinase		L05148.1	nt 189-1019	19
SYK	Tyr kinase		L28824.1	nt 121-116	2
PLCg1	Tyr kinase		4505868	nt 1485-235	4
A-raf-1	Ser/Thr kinase		X04790.1	nt 191-651	5
PAG	transmembrane adaptor		AF240634.1	nt 1-644	1
SHP/PTP1C	protein-tyrosine phosphatase		X62055 1	nt 746-1672	1
CSK	tyrosine kinase		4758077	nt 405-726	1
nucleolin (NCL)_	RNA binding		4885510	nt 3979-382	1
SLAP	Src-like adaptor protein		HSA238592	nt 3979-382	1
enzymes and receptors					
TRAC1	RING finger Ub ligase			1	19
PAK2	p21-activated kinase 2		4505598	nt 9-380	1
PAK2 (longer hit)	p21-activated kinase 2		4505598	nt 1-711	20
TCPTP/PTPN2	protein-tyrosine phosphatase		4506290	nt 3-1169	4
EDG1(hit2)	GPCR		XM_001499	nt 1-1192	1
EDG1(long or hit 3)	GPCR		XM_001499	nt 1-1288	1
IL10Ra	receptor		4504632	nt 750-141	1
integrin a2	receptor		6006008	nt 3390-395	2
Enolase 1a	phosphopyruvate hydratase		4503570	nt 797-1468	9
PRSM1	metallopeptidase		U58048.1	nt 1682-572	1
CLN2	protease		5597012	nt 1317-362	1
P2X5b	ionotropic ATP receptor		AF070573.1	nt 1212-70	1
6-PFKL	liver 6-phosphofructokinase		X16914.1	nt 260-1	1
DUSP1	dual specificity phosphatase		7108342	nt 1065-13	1
KIAA0251	Pyridoxal-dependent decarboxylase		D87438	nt 2098-237	2

FIG. 7

adaptors and transcription factors			
GG2-1	TNF-induced protein	AF070671.1	full-length	2
Grb7	adaptor	4885354	nt 1487-213	2
SH2-B	adaptor	AF227968.1	nt 1481-205	1
STAT1	signal transducing transcription factor	6274551	nt 113-737	1
TCF19	PHD finger transcriptional factor	11419132	nt 950-578	1
HFB101S	KRAB zinc finger protein	7706506	nt 562-372	1
RERE	transcriptional factor	AB036737	nt 1551-183	3
sudD	Ser/Thr rich	4507298	nt 15-506	1
Ku 70	DNA-PKc subunit	4503840	nt 1085-170	1
SCAMP2	secretory carrier membrane protein	AF005038.2	nt 25-863	1
Fibulin-5	EGF-like protein	5453649	nt 1024-827	1
KIAA1228	C2 domain (Ca2+ or IP binding)	AB033054	nt 1080-180	1
EST from clone 2108068	LPP20 lipoprotein precursor	AL357532.1	novel isofor	4
cytoskeleton			
vimentin	intermediate filament	4507894	nt 25-496	1
filamin A, alpha (FLNA)	actin-binding protein-280	4503744	nt 878-233	1
centraclin alpha(ACTR1A	A	5031568	nt 1481-825	1
moesin (MSN)	moesin (MSN)	4505256	nt 6-326	1
others			
TIMP3	inhibitor of metalloproteinase	9257248	nt 1225-140	1
RNH	ribonuclease/angiogenin inhibitor	4506564	nt 339-1073	1
HBA1	hemoglobin, alpha 1 and 2		partial	1
HBA1	hemoglobin, alpha 1 and 2		partial	1
TCRg	Ig superfamily		partial	2
MHCI	Ig superfamily		partial	1
MHCI	Ig superfamily		partial	2
IgG heavy chain	Ig superfamily		partial	7
Ig light chain	Ig superfamily		partial	1
IgG kappa	Ig superfamily		partial	4

FIG. 7 (CONTINUED)